

General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.
- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.
- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.
- This document is paginated as submitted by the original source.
- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

UNCLASSIFIED

AD 425 944

Processed by ...

CLEARINGHOUSE

FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION

of the

U.S. DEPARTMENT OF COMMERCE

N65-25949

FACILITY FORM 802

(ACCESSION NUMBER)

103

(PAGES)

(THRU)

3

(CODE)

AD 425944

(NASA CR OR TMX OR AD NUMBER)

06

(CATEGORY)

CR 63403

for ...

DEFENSE DOC

DEFENSE

GPO PRICE \$

OTS PRICE(S) \$

Hard copy (HC) 4.00

Microfiche (MF) .75

UNCLASSIFIED

Acquisitioned Document
597

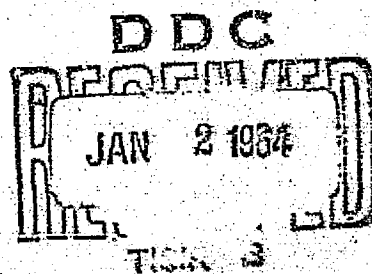
RECENT DEVELOPMENTS IN
PERTURBATION THEORY

by

Joseph O. Hirschfelder, W. Byers Brown
and Saul T. Epstein

University of Wisconsin
Theoretical Chemistry Institute
Madison, Wisconsin

To be published in
Advances in Quantum Chemistry
Volume I, 1964
Academic Press, New York



RECENT DEVELOPMENTS IN PERTURBATION THEORY*

by

Joseph O. Hirschfelder, W. Byers Brown

and Saul T. Epstein

University of Wisconsin, Theoretical Chemistry Institute
Madison, Wisconsin

ABSTRACT

25949

The techniques involved in some of the recent developments in low order perturbation theory are discussed in detail. These include: the close two-way relationship between the perturbation solutions and the variational principles; the existence of solutions to the perturbation equations and their convergence; the explicit and variational solutions of the first order perturbation equation; double perturbations and the calculation of expectation values; the determination of second order properties (such as polarizability); estimation of off-diagonal and time-dependent properties; fast converging iteration procedures; simplified treatment of degenerate or almost degenerate perturbations; sum rule techniques and their application to intermolecular forces; and a short survey of the atomic $1/Z$ expansion and Hartree-Fock applications of perturbation procedures.

JS *Spitzer*

* This research was supported by the following grant: National Aeronautics and Space Administration Grant NsG-275-62(4180).

TABLE OF CONTENTS
RECENT DEVELOPMENTS IN PERTURBATION THEORY

I. Introduction	1
II. Rayleigh-Schrodinger Perturbation Theory for a Non-Degenerate State	15
III. Solution of Perturbation Equations	22
A. Explicit Solution	23
B. Variational Solution	35
C. Variational Principles for Lower Bounds to the Energy	41
IV. Expectation Values and Double Perturbation Theory	46
A. Double Perturbation Theory	49
B. Dalgarno's Interchange Theorem	53
C. Solution of Perturbation Equations	56
D. Expansion in Perturbed and Unperturbed Eigenfunctions	57
V. General Connections Between Perturbation Theory and the Variational Principle	62
A. Perturbation Theory of the Variational Principle	63
B. Special Theorems for Variational Wave Functions	68
VI. The FOPIM and FOP-VIM Fast Converging Iteration Procedures	76
A. FOPIM	77
B. FOP-VIM	81
VII. Degenerate or Almost Degenerate Energy Level Perturbations	84
A. DE-FOP-VIM	85
VIII. Off-Diagonal Matrix Elements	90
A. General Formulation	90

VIII. Off-Diagonal Matrix Elements Cont'd

B. Variation Principles	92
C. Separation of Equations for One-Electron Operator	93
IX. Time Dependent Perturbation Theory	97
A. General Formalism	97
B. Double Perturbation Theory	101
X. Sum Rules and the Calculations of Sums	107
A. General Techniques for Evaluating Sums	108
B. Sum Rules for Oscillator Strengths	111
C. Application of the Sum Rules	123
XI. Applications	132
A. The $1/Z$ Expansion and the Shielding Approximations for Atoms	132
B. Perturbation Theory and the Hartree-Fock Approximation for Atoms.	134
C. Energy Levels of Isolated Systems	136
D. Expectation Values	141
E. Time Dependent Perturbation Theory	142
XII. Existence and Convergence of the Perturbation Series	148
A. Mathematical Properties of H	148
B. Perturbation Series	152
Acknowledgement	163
Appendix A - Algebraic Manipulations for Obtaining Perturbation Energies	164
Appendix B - Expansion in Terms of Arbitrary Complete Set	167
Appendix C - The Use of Delta Functions in the Hamiltonian and Other Operators	169

Appendix D - Solution of Separable Perturbation Equations in Generalized Coordinates	171
Appendix E - Solution of Perturbation Equation by Spherical Harmonic Expansion	177
References	180

I. INTRODUCTION

In recent years there has been a great increase of interest in the application of perturbation theory to the fundamental problems of quantum chemistry. Perturbation theory is designed to deal systematically with the effects of small perturbations on physical systems when the effects of the perturbations are mathematically too difficult to calculate exactly, and the properties of the unperturbed system are known. It is therefore usually the appropriate tool for dealing with the effects of external fields, with the long-range interactions between atoms and molecules, and with small internal perturbations such as those that give rise to the fine and hyperfine structure of spectral lines. Perturbation theory has been used for these purposes since the inception of quantum mechanics, but it is only recently that it has been applied seriously to what may be called the fundamental "few body" problems: electron repulsion and the many-center coulomb field in molecules.

The new applications have been mainly to atoms. The reciprocal of the atomic number, $1/Z$, provides a natural perturbation parameter, since the electron-electron repulsion terms have the form $(1/Z) r_{12}^{-1}$ when the coordinates are expressed in units of $Z a_0$ and the rest of the Hamiltonian is independent of Z . The new applications may be divided into two groups. The first consists of calculations of energy levels, and is a natural outgrowth of Hylleraas' classic work on the $1/Z$ expansion for two-electron atoms. It includes the continuation of Hylleraas' calculation to higher orders by Scherr and collaborators, the work of Dalgarno and Linderberg on the $1/Z$ expansions within the

Hartree-Fock approximation for atoms, and explorations by Parr and others on the united atom (single-center) approach to molecules. The applications in the second group are to the calculation of expectation values and other properties of atoms and molecules, and are of much more recent origin. They include calculations begun by Sternheimer and developed by Dalgarno and by Schwartz on the polarizabilities and nuclear shielding constants of atoms, and recent attempts to extend these ideas to molecules.

There are two principal reasons for the success of these new applications. First, sufficient accuracy is frequently obtained from knowledge of a first order perturbed wave function. From such a function, the energy can be computed accurate through the third order and good values can be obtained for the expectation values of properties other than the energy. Second, a great advantage of perturbation theory is that the functional form of the perturbed wave function is shaped by the perturbation itself. This is in contrast with the usual methods where the choice of trial function is arbitrary. For example, Schwartz (1959a), starting with the simplest hydrogenic wave function, obtained a simple closed-form perturbation expression for the charge density of the helium atom, which agrees almost precisely with the charge density calculated with the use of Hylleraas' 8 parameter wave function.

The recent renewal of interest in the potentialities of perturbation theory has been sparked by three developments.

(a) It had long been recognized in the literature, both implicitly and explicitly, that the standard summation formulae of perturbation theory are formal solutions of certain inhomogeneous equations, whose

solution might be better accomplished in some other fashion. Moreover in many cases an examination of the equations revealed the possibility of solution in closed form. However, in spite of successes, such techniques were not pursued systematically and, curiously enough, are rarely mentioned in the usual text books. Today, impressed by the many fruitful applications made by Dalgarno, by Schwartz, and by others, one is looking more and more to the equations themselves rather than to their formal solutions. The methods used to obtain explicit solutions of the perturbation equations are reviewed in Chapter IIIA.

(b) Although closed solutions are possible in some problems, in particular for many one-electron atomic problems, this is not the case when the perturbation involves the two-particle electron repulsion potential $1/r_{12}$, or indeed for most problems. Faced with this situation, there has been increasing recognition of the fact, pointed out long ago by Hylleraas, that one can often get satisfactory variational approximations to the individual terms in the perturbation series. The methods used to obtain variational solutions of the perturbation equations are reviewed in Chapter IIIB.

(c) Whereas the total energy calculated by means of approximate wave functions is accurate up to the second order, most physical properties calculated with such functions are only accurate up to the first order. The first correction to approximate properties for badness of the wave function can be written in terms of accessible functions by means of what may be called Dalgarno's Interchange Theorem. This is discussed and proved in Chapter IV.

These developments, which have been mainly applied to atoms, are

far from having exhausted their usefulness, or their power to stimulate further work. Perturbation theory appears destined to play an increasingly important role in quantum chemistry, and it therefore seems timely to review the basic theory underlying the new developments in a thorough manner. The present article is not intended to provide a comprehensive account of all aspects of perturbation theory, such as those given in the excellent review articles of Dalgarno (1962) and Preuss (1962). Nor are all the important new developments discussed, such as the infinite order perturbation treatments associated with the names of Brueckner (1955), Lowdin (1962) and Primas (1961), or the partial summation technique of Kelly (1963). Instead, most of our discussion is restricted to the first few orders of perturbation theory and devoted to the recent developments which have been most successful, or to those that seem most promising.

The only form of perturbation theory we discuss in this article is the Rayleigh-Schrodinger (reviewed in Chapter II) which was developed by Lord Rayleigh for vibrating systems and introduced into quantum mechanics by Schrodinger.¹ In this treatment, a Hamiltonian H for a system is regarded as consisting of an unperturbed Hamiltonian H_0 and a perturbation operator

$$\lambda V = H - H_0$$

In the Rayleigh-Schrodinger perturbation theory, the eigenvalues $E(\lambda)$ and eigenfunctions $\Psi(\lambda)$ of H are expanded in power series in λ . The essential role of λ is to define the different orders of perturbation, and this is a unique feature of the Rayleigh-Schrodinger theory. In certain problems, λ has an obvious physical significance.

Otherwise, λ is usually assigned the value of unity. Thus the perturbation equations arise from the Schrodinger equation

$$(H - E) \Psi = 0$$

by equating the coefficient of every individual power of λ to zero. There is a very important consequence of this procedure which it is appropriate to mention here, namely, that any theorem which is true for the exact solutions of the Schrodinger equation must have an analogue for the solutions of each perturbation equation separately. The most interesting examples are the variational principle (discussed in Chapter V), the virial and hypervirial theorems, and the Hellmann-Feynman theorem.

The variational principle and perturbation theory are deeply intertwined. First, given a quantity accurate through a particular order of perturbation, a corresponding variational principle can be constructed. Second, given a variational principle, individual variation principles for the different orders of perturbation can be derived and provide a fruitful and practical source of approximate perturbation energies and approximate perturbed wave functions. A striking instance of the connection is the fact that one of the most important theorems in perturbation theory, namely that knowledge of the n -th order wave function suffices to determine the energy to order $(2n+1)$, is an immediate consequence of the variation principle, as shown in Chapter II. These and other aspects of the variation principle and variational approximations are discussed at length in Chapter V.

A familiar complaint about perturbation theory is that when the perturbation becomes large, the treatment ceases to give meaningful

results. The usual rule of thumb is that the energy shifts should be small compared with the spacings of the unperturbed levels. If the energy levels tend to cross, one expects difficulties and special methods must be used (see Chapter VII).

However, almost degeneracy is the only source of difficulties. Sometimes the energy levels and eigenfunctions are not analytic functions of the perturbation parameter and no solutions to the equations exist. In other cases, the series may only be asymptotic.

The questions of the existence, nature, and convergence of the perturbation series pose difficult mathematical problems. Nevertheless some very powerful theorems, apparently little known to chemists, have been proved by Rellich (1939 and 1940), Kato (1951a), Titchmarsh (1958) and others, which cover most of the applications of interest in quantum chemistry. A basic theorem, due originally to Rellich (1939), is that the Rayleigh-Schrodinger perturbation series for $E(\lambda)$ and $\Psi(\lambda)$ converge for sufficiently small λ if the unperturbed Hamiltonian is self-adjoint (Chapter XII) and if two constants a and b can be found such that $V\phi$ and $H_0\phi$ satisfy the inequality

$$\langle V\phi, V\phi \rangle \leq a \langle H_0\phi, H_0\phi \rangle + b \langle \phi, \phi \rangle \quad (1)$$

for all functions ϕ in the domain of H_0 . The most important application of this theorem has been made by Kato (1951a), who succeeded in proving that it is satisfied for any decomposition of the potential U of the non-relativistic Schrodinger Hamiltonian operator H for any atom, molecule or finite crystal, provided no new singularities stronger

than Coulomb poles are introduced. In particular the theorem is true in the following important cases:

(a) The perturbation V is the electron repulsion term in the potential U . This result justifies the $1/Z$ expansion and proves that it is absolutely convergent for large enough Z . Although Kato could only obtain the crude estimate $Z \approx 7.6$ for the radius of convergence of the two-electron atom, the calculations of Scherr and Knight (1963) indicated that the true value is as low as $Z \approx 0.8$.

(b) The perturbation V is the difference between the many-center coulomb potential terms in a molecule and single-center terms for the united atom nucleus. This application justifies the united atom (single-center) expansions.²

For some unbounded operators the theorem of Eq. (1) is not satisfied and the series does not converge, but is nevertheless valid in the asymptotic sense; that is, useful only up to a certain finite order. Kato (1951a) has derived very general conditions under which the perturbation series will at least be asymptotic, although unfortunately these are harder to verify than Eq. (1). An even less well-behaved class of perturbations, namely those which are unbounded from below, causes the discrete spectrum of H_0 to vanish and be replaced by a continuous spectrum. A typical example is the Stark effect for an atom where the discrete energy levels are converted to metastable levels. The physical situation is quite clear in such cases of "weak quantization" and the series gives sensible results for small R , but the mathematical justification is quite difficult. All these aspects of the theory of convergence are described in more detail in Chapter XII.

The starting point in perturbation theory is the division of the intractable Hamiltonian H into a simple unperturbed part H_0 , whose eigenfunctions are known, and a remainder V containing the awkward terms. In the actual treatment of atoms and molecules, however, we are given H and some function ψ which is an approximation to the true wave function Ψ . The choice of ψ is generally made on the basis of physical intuition, mathematical expediency, etc. It is desired to treat ψ as a zero order wave function and use perturbation theory to improve it. In order to apply the usual formalism of perturbation theory it is necessary to construct an unperturbed Hamiltonian H_0 corresponding to ψ . It is important to realize that this problem can always be solved formally, as pointed out by Makinson and Turner (1953) and by Sternheimer (1954). We write the perturbed Hamiltonian H in the form

$$H = T + U \quad (2)$$

where T is the kinetic energy operator and U is the potential energy, assumed to be a function of coordinates. We then define the operator H_0 by

$$H_0 = T + U_0 \quad (3)$$

where

$$U_0 = \epsilon - \frac{TV}{\psi} \quad (4)$$

and the value of ϵ can be chosen arbitrarily. Then it follows from Eqs. (3) and (4) that ψ is an eigenfunction of H_0 with eigenvalue ϵ .

so that

$$H_0 \psi = E \psi \quad (5)$$

The perturbation potential V is then defined by ($\lambda = 1$)

$$V = U - U_0 \quad (6)$$

The next step in obtaining an improved wave function is to solve the first order equation. Dalgarno and Lewis (1955) defined a function F such that the first order wave function is equal to $F\psi$. The first order equation then becomes (for one electron),

$$\nabla \cdot (\psi^2 \nabla F) = 2\psi(V - \langle V \rangle)\psi \quad (7)$$

In analogy with electrostatics, this is a Poisson-type equation for the "potential" F produced by a charge distribution $-(2\pi)^{-1}\psi(V - \langle V \rangle)\psi$ in a region of variable dielectric constant ψ^2 (Prager and Hirschfelder, 1963). When separable, it can be integrated by quadrature; in which case the $F\psi$ procedure has determined the integrating factor. When Eq. (6) is not separable, approximate solutions for F can be obtained by using the excellent variational procedures discussed in Chapter IIIB.

The calculation of physical properties of a system, other than the energy, is a problem of great importance and interest. Typical properties are of two kinds, both associated with an operator W , and frequently also with an external field \mathcal{M} ; they are considered in detail in Chapter IV. The simplest properties are straight expectation values, $\langle W \rangle$, such as a dipole moment, and can be regarded as first order perturbation energies for a perturbation $\mathcal{M}W$. The other kind can be

regarded as the second order perturbation energies for the perturbation μW , and have the form

$$\langle Q \rangle = \langle \chi, (W - \langle W \rangle) \psi \rangle$$

where χ is the first order wave function corresponding to the Hamiltonian $H + \mu W$; a typical example is the electric polarizability. If the only wave function ψ available is approximate, with an error $\psi - \psi$ of order λ , the calculated values of $\langle W \rangle$ and $\langle Q \rangle$ will also be in error by terms of order λ ; only the energy $\langle E \rangle$ has an error of order λ^2 . The problem of devising a practical way of calculating the leading corrections has been solved by Dalgarno and Stewart (1956), and by Schwartz (1959), and is discussed in Chapter IV. The simplest approach is to use a double perturbation procedure (Dalgarno and Stewart, 1958) based on the Hamiltonian

$$H = H_0 + \lambda V + \mu W$$

Since the calculation of the first order energy for a perturbation only requires the corresponding zeroeth order wave function, the first order λ -correction to the W -perturbed energies (properties associated with W) only require perturbation solutions of the singly perturbed Hamiltonian

$$H_1 = H_0 + \mu W$$

This is the point of Dalgarno's Interchange Theorem (Dalgarno and Stewart, 1956; 1958). Its importance lies in the fact that the "badness" of the initial wave function is due to neglect of two-particle electron

repulsion terms which therefore appear in \bar{V} and make the first order λ -equation impossible to solve explicitly. The operator W , on the other hand, is usually the sum of one-particle terms, and the corresponding λ -equations are often easy to solve explicitly. The λ -corrections to $\langle W \rangle$ and $\langle Q \rangle$ then involve simply an integration over V , which is usually straightforward.

Perturbation methods are particularly valuable for the calculation of intermolecular potentials. The great advantage is that the interaction energy is calculated directly rather than obtained as the difference between two large numbers. In this manner, Dalgarno and Lynn (1956 and 1957) determined the long-range forces between two hydrogen atoms in excellent agreement at $R = 4a_0$ with the precise variational calculations of Kolos and Roothaan (1960). Other perturbation calculations have been made by Dalgarno and Lewis (1955 and 1956), Dalgarno and Stewart (1956), and by Salem (1960 and 1962). Since the wave functions are not known accurately for the separated molecules, the calculations of intermolecular potentials require not only double, but triple perturbation techniques.

The following more recent developments are described in the later chapters of the review.

1. Iterative Perturbation Methods (Chapters VI and VII)

Frequently it is sufficient to obtain the first order perturbed wave function since this permits the calculation of the energy accurate through the third order. Indeed, from the Rayleigh-Schrodinger wave function through the n -th order, the energy can be determined through the $(2n+1)$ -th order. However, even faster convergence can be obtained

by an iterative procedure (FOPIM) such that, from the n -th iterated wave function, the energy can be calculated up to terms involving the 2^{n+1} power of the perturbation parameter.

A combination of perturbation and variational techniques can be used (DE-FOP-VIM) to solve problems involving degenerate or almost-degenerate states without special consideration of the order in which the degeneracy is broken up or worry about the crossing of the energies as the perturbation parameter increases.

2. Time Dependent Problems (Chapters VIII and IX)

These are of two kinds: Those in which the perturbation causes transitions, and those in which the stationary properties are modified by an oscillating field. The techniques developed for time independent problems can be applied to both types. Interchange theorems have been proved in both cases under fairly general conditions by one of us (STE) and are presented here for the first time.³ Chapter VIII is devoted to second order steady states properties, and Chapter IX to the calculation of off-diagonal elements.

3. Sum Rules (Chapter X)

Useful Sum Rules can frequently be obtained by comparing the explicit solutions to the perturbation equations with their equivalent spectral representations expressed in infinite series. The family of sums involving oscillator strengths is of particular interest since it has a wide variety of physical applications. Sum rules can now be applied to dispersive intermolecular forces, thanks to an ingenious mathematical trick which has been pointed out by Mavroyannis and

Stephen (1962). Because of its importance, this application is reviewed in considerable detail.

4. Applications (Chapter XI)

This chapter contains a list of references to the recent literature on applications of perturbation theory to quantum chemistry. An effort has been made to make the bibliography complete apart from references to polarizability, nuclear shielding constants, etc., which have recently been reviewed by Dalgarno (1963).

In proceeding from the formal theory to the practical applications many questions arise. For example, is it really necessary or desirable to use Hartree-Fock orbitals as the starting point for perturbation calculations (Szasz, 1963)? The success of the work of Dalgarno and Stewart (1960) using Slater type orbitals with screening constants adjusted to make the first order perturbation vanish, suggests that simpler orbitals may provide a more convenient starting point. The perturbation equations themselves mold the wave function into the proper form.

The idea of applying perturbation theory to improve the familiar orbital model of atoms and molecules has been greatly stimulated by Sinanoglu's (1961) discussions of the electron correlation problem. Sinanoglu points out that the Hartree-Fock potential correctly represents the behavior when the electrons are far apart. It follows that the perturbation potential contains only short-ranged electron-electron interaction terms when the unperturbed wave function is of the self-consistent field type. Furthermore, because of the Pauli principle,

it is improbable that more than two electrons will come sufficiently close together to simultaneously feel the effects of these short-ranged forces. This supports the idea that electrons are mainly correlated in pairs belonging to the same spatial orbitals. And thus, in agreement with the G. N. Lewis notions which underlie present-day chemistry, the wave functions for even complex molecules can be broken up into sets of two-electron functions (geminals). Thus, to a chemically useful approximation, perturbation theory need only cope with problems involving two electrons at a time.

In proceeding from the formal theory to the practical application many questions arise. For example, is it really necessary or desirable to use Hartree-Fock orbitals as the starting point for perturbation calculations (Stewart, 1960)? The success of the work of Nakatsu and Stewart (1960) using Slater type orbitals with screening constants adjusted to make the first order perturbation vanish, suggests that simpler orbitals may provide a more convenient starting point. The perturbation equations themselves mold the wave function into the proper form.

The idea of applying perturbation theory to improve the familiar orbital model of atoms and molecules has been greatly stimulated by Sinanoglu's (1961) discussions of the electron correlation problem. Sinanoglu points out that the Hartree-Fock potential correctly represents the behavior when the electrons are far apart. It follows that the perturbation potential contains only short-ranged electron-electron interaction terms when the unperturbed wave function is of the self-consistent field type. Furthermore, because of the Pauli principle,

FOOTNOTES TO CHAPTER I

1. The Brillouin-Wigner or Feenberg perturbation series converge faster than the Rayleigh-Schrodinger, but they require that the energy be determined in an iterative manner.
2. Note that it does not justify expansions in powers of internuclear distances. These are, indeed, non-analytic in the simplest case of a one-electron diatomic molecule, where in fact a term in $R^5 \log R$ occurs (see Chapter XI).
3. The Interchange theorem for transitions has been proved independently using a variational method by Borowitz (private communication).

II. RAYLEIGH-SCHRODINGER PERTURBATION THEORY FOR A NON-DEGENERATE STATE

In order to lay the foundation for the new developments, let us review the Rayleigh-Schrodinger perturbation theory for a non-degenerate state. The unperturbed Hamiltonian H_0 has the complete set of normalized eigenfunctions ψ_j with the corresponding energies ϵ_j . The perturbed Hamiltonian $H = H_0 + \lambda V$ has the eigenfunctions Ψ_j with the corresponding energies E_j . We fix our attention on that perturbed state which in the limit as λ approaches zero has $\Psi_q = \psi_q$ and $E_q = \epsilon_q$ where ϵ_q is not degenerate.

The basic assumption of the Rayleigh-Schrodinger theory is that Ψ and E may be expanded in power series in the perturbation parameter:

$$\Psi = \sum_{n=0}^{\infty} \lambda^n \psi^{(n)} \quad \text{and} \quad E = \sum_{n=0}^{\infty} \lambda^n \epsilon^{(n)} \quad (1)$$

In order to keep the notation from becoming too clumsy, the subscript "q" is omitted except where it is necessary to avoid confusion between different states of the system. The power series (1) are substituted into the perturbed Schrodinger equation. Since the resulting equation must be true for all sufficiently small values of λ , the coefficient of each power of λ in the equation must be equal to zero. This leads to the family of perturbation equations. Clearly, since the state "q" is non-degenerate, $\psi^{(0)} = \psi$ and $\epsilon^{(0)} = \epsilon$, so that the zeroth order equation is

$$H_0 \psi = \epsilon \psi \quad (2)$$

where $H_0' = H_0 - \epsilon$. The first order equation is

$$H_0' \psi^{(1)} + V' \psi = 0 \quad (3)$$

where $V' = V - \epsilon^{(1)}$. And the n -th order relation is

$$H_0' \psi^{(n)} + V' \psi^{(n-1)} = \sum_{k=2}^n \epsilon^{(k)} \psi^{(n-k)}, \quad n = 2, 3, \dots \quad (4)$$

Normalization conditions are required to complete the specification of the perturbed functions.

We require that the exact wave function Ψ be normalized to unity.¹ Substituting the power series Eq. (2) into $\langle \Psi, \Psi \rangle = 1$ and requiring that the coefficient of each power of λ in this equation be individually zero, we obtain the normalization conditions for the various orders of perturbed wave functions,

$$\sum_{k=0}^n (j, n-k) = 0, \quad n = 1, 2, \dots \quad (5)$$

where $(j, k) = \langle \psi^{(k)}, \psi^{(j)} \rangle$. For $n = 1$, this gives $(0, 1) + (1, 0) = 0$, which for real functions is the familiar orthogonality condition $(0, 1) = 0$.

Expressions for the perturbation energies can be obtained by taking the scalar product of Eqs. (3) and (4) with ψ . Thus

$$\epsilon^{(1)} = \langle \psi, V \psi \rangle \quad (6)$$

$$\epsilon^{(2)} = \langle \psi, V \psi^{(1)} \rangle = \langle \psi, V' \psi^{(1)} \rangle \quad (7)$$

$$\epsilon^{(n)} = \langle \psi, V' \psi^{(n-1)} \rangle = \sum_{k=1}^{n-2} (0, k) \epsilon^{(n-k)}, \quad n = 3, 4, \dots \quad (8)$$

In deriving Eqs. (7) and (8), we have assumed that H_0 is Hermitian.

If V is also Hermitian, additional relations can be obtained for the

higher order perturbation energies by carrying out a sequence of algebraic manipulations of the perturbation equations, as explained in Appendix A. The most important of these relations are

$$\epsilon^{(3)} = \langle \psi^{(1)}, V' \psi^{(1)} \rangle \quad (9)$$

$$\epsilon^{(4)} = \langle \psi^{(1)}, V' \psi^{(2)} \rangle - \epsilon^{(2)}(2,0) \quad , \quad (10)$$

$$\epsilon^{(5)} = \langle \psi^{(2)}, V' \psi^{(2)} \rangle - \epsilon^{(2)}\{(1,2) + (2,1)\} \quad , \quad (11)$$

$$\epsilon^{(2n)} = \langle \psi^{(n-1)}, V' \psi^{(n)} \rangle - \sum_{k=2}^n \epsilon^{(k)} \sum_{j=0}^{k-1} (n+j-k, n-j) \quad , \quad n = 2, 3, \dots \quad (12a)$$

$$\epsilon^{(2n+1)} = \langle \psi^{(n)}, V' \psi^{(n)} \rangle - \sum_{k=2}^n \epsilon^{(k)} \sum_{j=0}^{k-1} (n+1+j-k, n-j) \quad , \quad n = 2, 3, \dots \quad (12b)$$

Eqs. (12) demonstrate the well-known theorem that the perturbation energy to the $(2n+1)$ -st order can be obtained from knowledge of the wave function perturbed through the n -th order. Hylleraas (1930) first showed that the energy through $\epsilon^{(3)}$ is determined by the wave function through $\psi^{(1)}$. Later, Dalgarno and Stewart (1956b), Dupont-Bourdelet, Tillieu, and Guy (1960), and Sinanoglu (1961a) made explicit demonstrations of the general theorem. Actually, this theorem is an immediate consequence of the variational principle given in Chapt. V (Wigner, 1935; Silverman, 1952; and Sinanoglu, 1961a). From the variational principle we know that if the trial wave function is in error by order η , the energy is in error by order η^2 . Putting $\eta = \lambda^{n+1}$ proves the theorem.

We shall have need for the explicit energy expectation

$\mathcal{E}(n) = \langle \Psi(n), H \Psi(n) \rangle$ corresponding to $\Psi(n)$, the perturbed wave function truncated after the n -th order term and normalized to unity,

$$\Psi(n) = \sum_{k=0}^n \lambda^k \psi^{(k)} / D(n) \quad (13)$$

Here the $D(n)$ is the normalization constant

$$D(n) = \left[1 + \lambda^{2n+1} \sum_{k=1}^n \lambda^{k-1} \sum_{j=0}^{n-k} (n-j, j+k) \right]^{-\frac{1}{2}} \quad (14)$$

The energy expectation is $\mathcal{E}(n) = \langle \Psi(n), H \Psi(n) \rangle$. Thus, we find

$$\mathcal{E}(0) = \epsilon + \lambda \epsilon^{(1)} \quad (15)$$

and

$$\mathcal{E}(1) = \epsilon + \lambda \epsilon^{(1)} + \frac{\lambda^2 \epsilon^{(2)} + \lambda^3 \epsilon^{(3)}}{1 + \lambda^2 (1,1)} \quad (16)$$

The results for larger values of n are given in Appendix A. In general, the $\mathcal{E}(n)$ are accurate through the order of λ^{2n+1} and provide a sequence of upper bounds to the exact energy, to which they converge.

Expansions in Unperturbed Wave Functions

In the usual presentation of Rayleigh-Schrodinger perturbation theory, the perturbed wave functions are expressed in terms of a spectral distribution of a complete function set. The formulae are given in Appendix B for a general function set which has no relation to the unperturbed problem (with the possible exception of one member of the set). Considerable simplifications are obtained if the function set is the complete set of unperturbed wave functions. Then (Dalgarno, 1961) using the notation $V_{jk} = \langle \psi_j, V \psi_k \rangle$ and $V'_{jk} = \langle \psi_j, V' \psi_k \rangle$,

$$\psi^{(1)} = \sum_j' \frac{V_{jq} \psi_j}{\epsilon_q - \epsilon_j} \quad (17)$$

$$\psi^{(2)} = \sum_j' \sum_k' \frac{V_{qj} V_{jk} \psi_k}{(\epsilon_q - \epsilon_j)(\epsilon_q - \epsilon_k)} - \frac{1}{2} \sum_j' \frac{V_{qj} V_{jq}}{(\epsilon_q - \epsilon_j)^2} \psi_q \quad (18)$$

$$\begin{aligned} \psi^{(3)} = & \sum_j' \sum_k' \sum_l' \frac{V_{qj} V_{jk} V_{kl} \psi_l}{(\epsilon_q - \epsilon_j)(\epsilon_q - \epsilon_k)(\epsilon_q - \epsilon_l)} \\ & - \frac{1}{2} \sum_j' \sum_k' \frac{V_{qj} V_{jq} V_{qk} \psi_k}{(\epsilon_q - \epsilon_j)^2 (\epsilon_q - \epsilon_k)} - \epsilon_q^{(2)} \sum_j' \frac{V_{qj} \psi_j}{(\epsilon_q - \epsilon_j)^2} \\ & - \sum_j' \sum_k' \frac{V_{qj} V_{jk} V_{kq} \psi_q}{(\epsilon_q - \epsilon_j)^2 (\epsilon_q - \epsilon_k)} \end{aligned} \quad (19)$$

Here the primed sigmas indicate the sum over all of the discrete states plus the integral over the continuum but omitting the state (or states) with energy ϵ_q .

The corresponding expressions for the second and third order perturbation energies are

$$\epsilon^{(2)} = \sum_k' \frac{V_{qk} V_{kq}}{\epsilon_q - \epsilon_k} \quad (20)$$

and

$$\epsilon^{(3)} = \sum_j' \sum_k' \frac{V_{qj} V_{jk} V_{kq}}{(\epsilon_q - \epsilon_j)(\epsilon_q - \epsilon_k)} \quad (21)$$

As Brueckner (1955), Huby (1961), Primas (1961), and Lowdin (1962) have shown, compact explicit formulae for the higher order perturbation energies and wave functions may be written in terms of the operator

$$Q = \sum_j \frac{|\psi_j\rangle\langle\psi_j|}{(\epsilon_q - \epsilon_j)} \quad (22)$$

Thus for example the energies are :

$$\epsilon^{(4)} = \langle \psi, VQ [V'QV' - \epsilon^{(2)}] QV \rangle \quad (23)$$

$$\epsilon^{(5)} = \langle \psi, VQ [V'QV'QV' - \epsilon^{(3)} - \epsilon^{(2)}\{V'Q + QV'\}] VQ \rangle \quad (24)$$

...

FOOTNOTES TO CHAPTER II

1. There is another type of normalization which is commonly used.

Instead of requiring that $\langle \Psi, \Psi \rangle = 1$, the perturbed components of the wave function are taken to be orthogonal to the original function, $(0, n) = 0$ for $n = 1, 2, \dots$. This makes $\epsilon^{(n)} = \langle \psi, V\psi^{(n-1)} \rangle$. Or, according to F. Dupont-Bourdelet, J. Tillieu, and J. Guy [J. Phys. Radium 21, 776 (1960)] :

$$\epsilon^{(2n)} = \langle \psi^{(n-1)}, V\psi^{(n)} \rangle - \sum_{k=1}^{n-1} \sum_{j=n-k}^{2n-k-1} \epsilon^{(j)}(k, 2n-k-j)$$

$$\epsilon^{(2n+1)} = \langle \psi^{(n)}, V\psi^{(n)} \rangle - \sum_{k=1}^n \sum_{j=n+1-k}^{2n-k} \epsilon^{(j)}(k, 2n+1-k-j)$$

than that of the original perturbed equation because they are infinite. General differential equations for which general techniques are available whereas the original equation is an eigenvalue equation. This is

III. SOLUTION OF PERTURBATION EQUATIONS

In this chapter methods for solving the perturbation equations of Rayleigh-Schrodinger perturbation theory are discussed. The n -th order equation (2.4) for $\psi^{(n)}$ involves all the wave functions of lower order, so that the equations must be solved step-wise starting with the first order equation (2.3), substituting its solution into the second, solving the second order equation, and so on. The same is true for the equations of double perturbation theory discussed in Chapter IV. The basic problem is therefore to solve a first order perturbation equation of the type

$$(H_0 - \epsilon)\psi^{(1)} + (V - \epsilon^{(1)})\psi = 0 \quad (1)$$

where H_0 is the unperturbed Hamiltonian, ψ and ϵ are the unperturbed eigenfunction and energy (assumed non-degenerate) of the state under consideration, $\psi^{(1)}$ and $\epsilon^{(1)}$ are the first order eigenfunction and energy, and V is the perturbation operator. The solution of this equation makes it possible to evaluate the second and third order energies by means of Eqs. (2.7) and (2.9). However the recent developments mentioned in the introduction have been more concerned with using the solution of this type of equation to obtain the first order corrections to expectation values. In this case the perturbation is the operator W whose expectation value is sought.

The solution of perturbation equations is usually much easier than that of the original perturbed equation because they are inhomogeneous differential equations for which general techniques are available, whereas the original equation is an eigenvalue equation. This is

especially true in the most usual situation when H_0 is a one-electron Hamiltonian or the sum of one-electron Hamiltonians. Unfortunately, however, in the case of the perturbation of greatest importance, namely the electron repulsion terms $1/r_{12}$, analytic solutions have not been found and their discovery is no longer a matter for hope. The same is true of other two-electron perturbations such as the dipole-dipole and higher multipole interaction terms occurring in the treatment of long-range forces. The first order equations involving two-electron terms in V must therefore be treated by alternative approximate methods such as the variation method discussed in section B of this Chapter or the familiar method of expansion in a complete set as discussed in Chapter II and Appendix B. The explicit solution of perturbation equations in closed form is therefore effectively limited to perturbations which only involve one-electron operators or the sum of one-electron operators. Fortunately most properties of interest other than the energy are associated with operators W of this kind.

A. Explicit Solution

There is a well known device in the treatment of inhomogeneous differential equations (see Morse and Feshbach, 1953) which is used to simplify (1), and which has been employed very effectively by Dalgarno and his colleagues (Dalgarno and Lewis, 1955). This is to put

$$\psi^{(1)} = F\psi, \quad (2)$$

where F is a scalar function of the coordinates to be determined, and to note that

$$(H_0 - \epsilon)F\psi = (H_0 F - FH_0)\psi$$

so that (1) may be written

$$[H_0, F]\psi + V'\psi = 0, \quad (3)$$

where the square brackets denote a commutator and $V' = V - \epsilon^{(1)}$.

This form already has the advantage that it is usually unnecessary to know the explicit form of the unperturbed Hamiltonian H_0 , since

$$[H_0, F] = [H, F]$$

as long as the perturbation $H - H_0$ is a scalar, and V' can be replaced by $H - \epsilon - \epsilon^{(1)}$ (taking $\lambda = 1$). Equation (3) can then be written in the form

$$[H, F]\psi + (H - \epsilon - \epsilon^{(1)})\psi = 0 \quad (4)$$

To proceed further it is necessary to specify the form of the Hamiltonian. For the many electron Hamiltonian (atomic units: energy $\sim e^2/a_0$ and length $\sim a_0$),

$$H_0 = -\frac{1}{2} \sum_i \nabla_i^2 + U_0, \quad (5)$$

where the unperturbed potential energy U_0 is a scalar, so that (3) becomes

$$\sum_i [\nabla_i^2, F]\psi = 2V'\psi \quad (6)$$

If this equation is multiplied by ψ^* and its complex conjugate by ψ , the sum of the two equations can be rearranged to give

$$\sum_i \nabla_i (\psi^* \nabla_i F) = \psi^* \nabla^2 F + \psi \nabla^2 \psi^* \quad (7)$$

where we have assumed we can choose F to be real.¹ For simplicity of presentation we shall only consider the case when ψ is real so that (7) becomes

$$\sum_i \nabla_i (\psi^2 \nabla_i F) = 2\psi \nabla^2 F \quad (8)$$

The boundary conditions on the perturbed wave function, and therefore on all $\psi^{(n)}$, are the same as those on the unperturbed eigenfunction ψ . Thus on the boundaries at infinity the conditions on F are:

$$F\psi = 0, \quad \psi \nabla^2 F = 0 \quad (9)$$

This guarantees that the integral of the right-hand side of (8) vanishes.

The solution F of (8) is of course arbitrary to the extent of an additive constant, which may be chosen so that the orthogonality condition (2.5) is satisfied; that is

$$\langle \psi, F\psi \rangle = 0 \quad (10)$$

In terms of F the second order energy, $\epsilon^{(2)}$ given by (2.7) is

$$\epsilon^{(2)} = \langle \psi, \nabla^2 F \psi \rangle \quad (11)$$

By substituting from (8) and using Green's theorem we get

$$\begin{aligned} \epsilon^{(2)} &= \frac{1}{2} \sum_i \langle \nabla_i (\psi^2 \nabla_i F) \rangle \\ &= -\frac{1}{2} \langle \psi, \sum_i (\nabla_i F)^2 \psi \rangle \end{aligned} \quad (12)$$

since the surface integral vanishes because of the boundary condition (9). Thus the second order energy need only involve $\nabla_1 F$ in the integrand rather than F itself, and in principle only a single integration of (8) is needed. In terms of F the third order energy given by (2.9) becomes

$$\epsilon^{(3)} = \langle \psi, FV'FV \rangle \quad (13)$$

$$= \frac{1}{2} \sum_i \langle F^2 \nabla_1 (\psi^2 \nabla_1 F) \rangle$$

$$= - \sum_i \langle \psi, F(\nabla_1 F)^2 \psi \rangle \quad (14)$$

The integral for the third order energy thus involves F explicitly and the forms (13) and (14) are only valid if F satisfies the orthogonality condition (10).

1. Reduction of the Many Electron Equation

For completeness let us indicate how the many electron problem may be reduced to one-electron equations in the simplest case. This is when the total Hamiltonian H is the sum of one-electron Hamiltonians

$$H = \sum_i \{h(i) + \lambda v(i)\} \quad (15)$$

The only significant remark to be made is that it is best to separate the many electron equation into the one-electron equations

$$(h + \lambda v)\phi_i = \epsilon_i \phi_i \quad (16)$$

before considering the perturbation problem. The alternative is to attempt to reduce the total first order equation, and is neither as simple nor as accurate. The one-electron first order perturbation equations derived from (16) are

$$(h - \epsilon_i) \phi_i^{(1)} + (v - \epsilon_i^{(1)}) \phi_i^{(0)} = 0, \quad (17)$$

and the total wave function correct to the first order in λ for each electron is now

$$\Psi = \mathcal{A} \prod_i \{ \phi_i^{(0)} + \lambda \phi_i^{(1)} \}, \quad (18)$$

where \mathcal{A} is the anti-symmetrizer and the ϕ 's include spin factors. This procedure leads to simple sums over the electrons for all the perturbation energies; for example

$$\begin{aligned} \epsilon^{(1)} &= \sum_i \epsilon_i^{(1)} = \sum_i \langle \phi_i^{(0)}, v \phi_i^{(0)} \rangle, \\ \epsilon^{(2)} &= \sum_i \epsilon_i^{(2)} = \sum_i \langle \phi_i^{(1)}, (v - \epsilon_i^{(1)}) \phi_i^{(0)} \rangle \end{aligned} \quad (19)$$

2. One-Dimensional Equations

As an introduction to the solution of equation (8) let us consider the simplest case of a one-dimensional system described by a variable x confined to the interval (a,b) . Equation (8) may then be integrated directly to give

$$\psi^2 \frac{dF}{dx} = M(x) \quad (20)$$

where

$$M(x) = 2 \int_a^x \psi(\xi) \psi'(\xi) \psi(\xi) d\xi, \quad (21)$$

and the constant of integration vanishes because of the boundary condition (9).

(a) Ground State. For the ground state in which the unperturbed wave function ψ does not possess any nodes inside the interval (a, b) , (20) may be integrated immediately to give

$$F(x) = F(a) + \int_a^x M(\xi) \psi^{-2}(\xi) d\xi \quad (22)$$

as pointed out by Young and March (1958). The one-dimensional equations for $\epsilon^{(2)}$ and $\epsilon^{(3)}$ corresponding to (11) and (13) can be written in terms of M as

$$\epsilon^{(2)} = -\frac{1}{2} \int_a^b \left[\frac{M(x)}{\psi(x)} \right]^2 dx \quad (23)$$

$$\epsilon^{(3)} = - \int_a^b dx \left\{ \left[\frac{M(x)}{\psi(x)} \right]^2 + 2 \epsilon^{(2)} \psi^2(x) \right\} \int_a^x d\xi \frac{M(\xi)}{\psi^2(\xi)} \quad (24)$$

(b) Excited States. Consider now the case in which $\psi^{(x)}$ is the n -th unperturbed eigenfunction with n simple zeros at a_1, a_2, \dots, a_n in the interval (a, b) . This case has been treated by Brown and Hirschfelder (1963). If the first order wave function $\psi^{(1)}$ is well behaved the function F will have simple poles at a_1, a_2, \dots, a_n . The direct integration of (20) to yield F presents difficulties because $M(x)$ does not in general vanish at the nodes of ψ , and for the same reason (23) and (24) are not valid for excited states. To avoid the singularities in these equations let us assume that

ψ, F and M are analytic functions in the vicinity of the real interval (a, b) . It may be shown that equations (22) and (23) can then be replaced by

$$F(x) = F(a) + \oint_a^x M(z) \psi^{-2}(z) dz \quad (25)$$

and

$$\epsilon^{(2)} = -\frac{1}{2} \oint_a^b \left[\frac{M(z)}{\psi(z)} \right]^2 dz \quad (26)$$

where the contours C in the z -plane do not pass through a_1, a_2, \dots, a_n . A similar equation can be written for $\epsilon^{(3)}$ in place of (24).

An alternative method for dealing with excited states, more suited to numerical work, is to remove the singularities by subtracting out the poles of F . This may be accomplished by defining a new function $\Phi(x)$ by

$$F = \Phi - \sum_{j=1}^n \frac{C_j}{x - a_j} \quad (27)$$

and choosing the coefficients C_j so that Φ is analytic at the points a_j . By substituting (27) into the differential equation (20) we get

$$\frac{d\Phi}{dx} = \frac{M(x)}{\psi^2(x)} - \sum_{j=1}^n \frac{C_j}{(x - a_j)^2} \quad (28)$$

The coefficients C_j are therefore given by

$$C_j = M(a_j) / [\psi'(a_j)]^2 \quad (j = 1, 2, \dots, n) \quad (29)$$

Equation (28) may now be integrated along the real axis to give

$$\Phi(x) = \Phi(a) + \int_a^x \left[\frac{M(\xi) - \Lambda(\xi)}{\psi^2(\xi)} \right] d\xi \quad (30)$$

where

$$\Lambda(x) = \psi^2(x) \sum_{j=1}^n \frac{C_j}{(x - a_j)^2} \quad (31)$$

and the second order energy is given by

$$\epsilon^{(2)} = -\frac{1}{2} \int_a^b M \left[\frac{M - \Lambda}{\psi^2} \right] dx - \frac{1}{2} \sum_{j=1}^n C_j \int_a^b \frac{\psi V' \psi}{x - a_j} dx. \quad (32)$$

(c) Separable Equations. Direct quadrature, as in the one-dimensional cartesian case above, is also possible when the original Schrodinger equation is separable in curvilinear coordinates. In this case, as in the reduction of the many electron equation, it is advantageous to separate the exact Schrodinger equation into ordinary differential equations first, and then to expand into the different orders of perturbation. This avoids spurious cross terms in the energy of the second and higher orders. If the perturbation V appears in two or more of the separated equations the method is really a multiple perturbation procedure for independent perturbations, so that cross terms are rigorously absent. Details of the separation and equations are given in appendix D. The most important chemical application is to the two-center problem in molecules which is separable in confocal elliptic coordinates (see Chapter XI).

3. Separation by Partial Expansion

The perturbation equations can be solved explicitly in a much wider class of cases than the strictly separable one of the previous

sub-section. When a one-electron equation of the form (1) is not separable as it stands, the first step towards finding a solution is usually to expand the perturbation V and the wave function $\psi^{(1)}$ in terms of a complete orthogonal set in one or two of the coordinates. To illustrate the method consider first the important case of a spherically symmetric unperturbed wave function $\psi(r)$ perturbed by a potential $V(r, \theta, \phi)$ which may be expanded in spherical harmonics $Y_{\ell m}$:

$$V = \sum_{\ell, m} V_{\ell m}(r) Y_{\ell m}(\theta, \phi) \quad (33)$$

F may be similarly expanded to give

$$F = \sum_{\ell, m} F_{\ell m}(r) Y_{\ell m}(\theta, \phi) \quad (34)$$

By substituting these expansions into (8) and separating the harmonic components we get

$$\frac{d}{dr} \left(r^2 \psi^2 \frac{dF_{\ell m}}{dr} \right) - \ell(\ell+1) \psi^2 F_{\ell m} = 2r^2 \psi V_{\ell m} \psi, \quad (m = -\ell, \dots, \ell; \ell = 0, 1, 2, \dots). \quad (35)$$

This is a set of ordinary linear inhomogeneous second order differential equations which, although they are not directly integrable by quadrature for $\ell \neq 0$, may be solved either by inspection or by standard procedures.

From the formal standpoint the key to solving an equation of the type (35) is a solution of the corresponding homogeneous equation, which may be written

$$\frac{d^2 K}{dr^2} + 2 \frac{d \ln(r\psi)}{dr} \frac{dK}{dr} - \frac{\ell(\ell+1)}{r^2} K = 0 \quad (36)$$

If such a solution K is known we repeat the device (2) used in simplifying the original perturbation equation and put

$$F_{\ell m} = Kf \quad (37)$$

By substituting into (35), multiplying by K and rearranging, we get

$$\frac{d}{dr} \left(r^2 \psi^2 K^2 \frac{df}{dr} \right) = 2r^2 K \psi V'_{\ell m} \psi \quad (38)$$

The equation is now in the one-dimensional form discussed in sub-section 2(a) of this Chapter and may be integrated immediately to give²

$$r^2 \psi^2 K^2 \frac{df}{dr} = M_{\ell m}(r) \quad (39)$$

where

$$M_{\ell m}(r) = 2 \int_0^r s^2 K \psi V'_{\ell m} \psi ds \quad (40)$$

For the ground state Eq. (39) may be integrated again directly. For excited states the unperturbed wave function ψ will possess nodes, and so also may the solution K of the homogeneous Eq. (36). This situation has been discussed in sub-section 2(b) of this Chapter. The singularities of f may be either avoided by integrating along a contour C in the complex plane which avoids the zeros of ψ and K , or the poles may be subtracted out. For reasons of space we shall give only the complex integral form of F which is

$$F_{\ell m}(r) = K(r) \left\{ f(0) + \oint_C^r \frac{M_{\ell m}}{s^2 \psi^2 K^2} ds \right\} \quad (41)$$

The second order energy can be written in a form analogous to (26)

$$\epsilon^{(2)} = -\frac{1}{2} \sum_{l,m} \int_0^\infty \left[\frac{M_{lm}}{3\psi K} \right]^2 ds, \quad (42)$$

and likewise the third order energy. In many cases, however, it may be easier and quicker to obtain a particular integral of the inhomogeneous equation than to solve the homogeneous one and integrate.

The case in which ψ is an orbital eigenfunction with non-zero angular momentum L ,

$$\psi = R_L(r) Y_{LM}(\theta, \phi) \quad (43)$$

is similar in principle but somewhat more complicated in general, and is therefore treated in Appendix E. However when the perturbation V is radial, that is for the component V'_{00} , the first order equation may be integrated directly to give

$$F(r) = F(c) + \int_0^r \left[M/s^2 R_L^2 \right] ds \quad (44)$$

where

$$M(r) = 2 \int_0^r s^2 R_L V'_{00} R_L ds \quad (45)$$

The second order energy can be written as usual either in the form (see Brown and Hirschfelder, 1963)

$$\epsilon^{(2)} = -\frac{1}{2} \int_0^\infty \left[M/s R_L \right]^2 ds \quad (46)$$

or

$$\epsilon^{(2)} = -\frac{1}{2} \int_0^\infty M \left[\frac{M - \Gamma}{r^2 R_L^2} \right] dr - \frac{1}{2} \sum_j C_j \int_0^\infty \frac{r^2 R_L V'_{00} R_L}{r - a_j} dr \quad (47)$$

where a_j are the zeros of R_L and

$$C_j = M(a_j) / [a_j R_L'(a_j)]^2 \quad (48)$$

$$T(r) = r^2 R_L^2 \sum_j \frac{C_j}{(r - a_j)^2} \quad (49)$$

4. Second and Higher Order Equations

So far we have only considered the first order equation. The second order perturbation equation corresponding to (1) can be obtained from (2.4) and is

$$(H_0 - \epsilon) \psi^{(2)} + V' \psi^{(1)} - \epsilon^{(2)} \psi = 0 \quad (50)$$

It has the same form as Eq. (1) except that the inhomogeneous term involves $\psi^{(1)}$ which must be known before $\psi^{(2)}$ can be found. To solve Eq. (50) we therefore put

$$\psi^{(2)} = G\psi \quad (51)$$

and suppose that $\psi^{(1)} = F\psi$ is known. Then by the manipulations described at the beginning of this section we get

$$\nabla \cdot (\psi^2 \nabla G) = 2\psi(V'F - \epsilon^{(2)})\psi \quad (52)$$

This equation is now of the same form as Eq. (8), and can be solved by the same techniques.

The greatest need for explicit forms of the second order perturbation functions is in order to correct the zeroth order approximation to a second order property; this is discussed Chapter IV. The search for explicit solutions of the higher order equations follows the same

pattern as that of the first and second order equations. However they are not usually of special significance.

B. Variational Solution

When the unperturbed equation is not separable, or when the partial expansion technique requires an infinite series which only converges slowly, it is nevertheless often possible to obtain a useful variational approximation to the perturbed wave functions. This is indeed the only effective method when the perturbation is due to electron repulsion.

The basis of the variational approach is discussed fully in Chapter V. There is a variational principle of a very general kind for each order of perturbation which may be derived from the Ritz variation principle for the total energy. In practice it is natural to choose a special form of these principles, of which the prototype is that for the first order wave function introduced by Hylleraas (1930).

1. Hylleraas Variation Principle for Ground State

Hylleraas pointed out that for the ground state of a system Eq. (1) is equivalent to the variation principle

$$\tilde{\epsilon}^{(2)} \geq \epsilon^{(2)} \quad (53)$$

where

$$\tilde{\epsilon}^{(2)} = \langle \tilde{\psi}^{(1)}, (H_0 - \epsilon) \tilde{\psi}^{(1)} \rangle + \langle \tilde{\psi}^{(1)}, V \psi \rangle + \langle \psi, V \tilde{\psi}^{(1)} \rangle \quad (54)$$

and $\tilde{\psi}^{(1)}$ is arbitrary; the equality in (53) only holds when the variation function $\tilde{\psi}^{(1)} = \psi^{(1)}$, the true first order wave function. The principle is easy to prove: by substituting $\tilde{\psi}^{(1)} = \psi^{(1)} + \delta \tilde{\psi}^{(1)}$ in (2) we get

$$\tilde{\epsilon}^{(2)} - \epsilon^{(2)} = 2\langle \delta\tilde{\psi}^{(1)}, [(H_0 - \epsilon)\tilde{\psi}^{(1)} + V\tilde{\psi}^{(1)}] \rangle + \langle \delta\tilde{\psi}^{(1)}, (H_0 - \epsilon)\delta\tilde{\psi}^{(1)} \rangle \quad (55)$$

The term in δ vanishes because of (1) and the term in δ^2 is positive because ϵ is the lowest eigenvalue of H_0 . Clearly the same argument also applies to excited states which are the lowest of each symmetry class of the total Hamiltonian $H = H_0 + \lambda V$.

If the first order equations cannot be solved analytically it may nevertheless be possible to get a good approximation to $\epsilon^{(2)}$ and $\tilde{\psi}^{(1)}$ by inserting a trial function $\tilde{\psi}^{(1)}$ in (54) and making $\tilde{\epsilon}^{(2)}$ stationary with respect to the variation parameters appearing in $\tilde{\psi}^{(1)}$. If the trial function is a linear combination of orthogonal functions ϕ_α then $\tilde{\epsilon}^{(2)}$ has a form similar to Eq. (B.6). This procedure was introduced by Hylleraas and has been used by him (1956; 1958) in approximating to the second order energy in the $1/Z$ expansion of two-electron atoms. This work on He-like systems has been continued to higher orders of perturbation of Scherr and Knight (1963), and is discussed in Chapter XI.

One of the most important applications of the variational principle is to determine an approximate $\psi^{(1)}$ in order to calculate the first order corrections to expectation values (see Chapter IV). In cases where the unperturbed wave function ψ is not known exactly, Eq. (54) may still be used to determine an approximate $\psi^{(1)}$ by making $\tilde{\epsilon}^{(2)}$ stationary, but in such a case $\tilde{\epsilon}^{(2)}$ is no longer a correct variational approximation to $\epsilon^{(2)}$. This approach has been used extensively in shielding calculations and is further discussed in Chapter V (see in particular footnote 8, where the functional $\tilde{\epsilon}^{(2)}$ is denoted by \tilde{J}_H).

(a) Dirichlet Form of Hylleraas Principle. If we put $\tilde{\psi}^{(1)} = \tilde{F}\psi$ as in Eq. (2) and regard the magnitude of \tilde{F} as a variable, we can minimize $\tilde{\epsilon}^{(2)}$ explicitly to get a form homogeneous in \tilde{F}

$$\tilde{\epsilon}^{(2)} = - \frac{\langle \psi, \nabla^2 \tilde{F} \psi \rangle^2}{\langle \psi, \tilde{F} (H_0 - \epsilon) \tilde{F} \psi \rangle} \quad (56)$$

The denominator in this expression can be written

$$\langle \psi, \tilde{F} (H_0 - \epsilon) \tilde{F} \psi \rangle = \langle \psi, \tilde{F} (H_0 \tilde{F} - \tilde{F} H_0) \psi \rangle,$$

and if H_0 has the form of Eq. (5) then

$$\tilde{F} (H_0 \tilde{F} - \tilde{F} H_0) \psi = \frac{1}{2} \sum_i (\psi \nabla_i^2 \tilde{F})^2 - \frac{1}{2} \sum_i \nabla_i \cdot (\tilde{F}^2 \nabla_i \tilde{F})$$

On integrating this expression we get

$$\langle \psi, \tilde{F} (H_0 \tilde{F} - \tilde{F} H_0) \psi \rangle = \frac{1}{2} \langle \psi, \sum_i (\nabla_i^2 \tilde{F})^2 \psi \rangle$$

provided that the surface integrals of $\tilde{F}^2 \nabla_i \tilde{F}$ vanish. Equation (56) can then be written

$$\tilde{\epsilon}^{(2)} = - 2 \frac{\langle \psi, \nabla^2 \tilde{F} \psi \rangle^2}{\langle \psi, \sum_i (\nabla_i^2 \tilde{F})^2 \psi \rangle} \geq \epsilon^{(2)} \quad (57)$$

It is clear from this form that the second order energy for the ground state is always negative.

Prager and Hirschfelder (1963) have recently pointed out that the first order Eq. (8) is analogous to a 3N-dimensional Poisson's equation in electrostatics, F playing the role of the electrostatic potential, ψ^2 that of a variable dielectric constant and $-\psi \nabla^2 \psi / 2\pi$ that of the charge density. The variation principle in the form (57) is well

known in electrostatics as the Dirichlet principle for the self energy of a charge distribution.

A particularly simple general type of variation function which has proved useful in connection with this formula when V is scalar is

$$\tilde{V} = V' \quad (58)$$

which leads to

$$\tilde{\epsilon}^{(2)} = -2 \frac{\langle \psi, (V')^2 \psi \rangle^2}{\langle \psi, \sum_i (\nabla_i V)^2 \psi \rangle} \geq \epsilon^{(2)} \quad (59)$$

Unfortunately the denominator diverges in the important case of electron repulsion so that $\tilde{\epsilon}^{(2)} = 0$.

2. Extension of Hylleraas Principle to Excited States

The usual extension of the Ritz variation principle to excited states requires knowledge of the exact perturbed wave functions of the lower states of the same symmetry. Approximate energies and wave functions for the excited states can of course be obtained in the process of minimizing the trial energy expression for the ground state, but in general these approximations are unbounded. The only case in which the approximate energy is an upper bound to the excited state energy is when it is the lowest state of a given symmetry type. An important advantage of the more restricted variation principle for the second order energy is that, as shown by Sinanoglu (1961b), it can be fairly easily modified to apply to an excited state if the unperturbed wave functions for the lower states are known.

Consider the variation $\delta \tilde{\epsilon}_q^{(2)}$ for the excited state q when

$\tilde{\epsilon}_q^{(2)}$ is stationary with respect to changes $\delta \tilde{\psi}_q^{(1)}$:

$$\delta \tilde{\epsilon}_q^{(2)} = \langle \delta \tilde{\psi}_q^{(1)}, (H_0 - \epsilon_q) \delta \tilde{\psi}_q^{(1)} \rangle \quad (60)$$

The reason this is not always positive is that the variations will in general contain components ψ_k ($k < q$) which give rise to terms $(\epsilon_k - \epsilon_q)$ which are negative. These components are in fact known exactly from perturbation theory, and the variational method suggested by Sinanoglu is to insert them in the first order trial wave function $\tilde{\psi}_q^{(1)}$ and to vary only that part of $\tilde{\psi}_q^{(1)}$ which is orthogonal to all the lower states. That is, we put

$$\tilde{\psi}_q^{(1)} = \sum_{k < q} \frac{V_{kq}}{\epsilon_q - \epsilon_k} \psi_k + \tilde{\chi}_q^{(1)} \quad (61)$$

where

$$\langle \psi_k, \tilde{\chi}_q^{(1)} \rangle = 0 \quad (k < q) \quad (62)$$

By substituting (61) into (54) we get

$$\tilde{\epsilon}_q^{(2)} = \sum_{k < q} \frac{|V_{qk}|^2}{\epsilon_q - \epsilon_k} + \langle \tilde{\chi}_q^{(1)}, (H_0 - \epsilon_q) \tilde{\chi}_q^{(1)} \rangle + 2 \langle \tilde{\chi}_q^{(1)}, V' \psi_q \rangle \quad (63)$$

where $V' = V - \epsilon_q^{(1)}$. Let $\chi_q^{(1)}$ be the exact part of $\psi_q^{(1)}$ orthogonal to ψ_k ($k < q$), and put $\tilde{\chi}_q^{(1)} = \chi_q^{(1)} + \delta \tilde{\chi}_q^{(1)}$. Then by the analogue of Eq. (55) we have

$$\tilde{\epsilon}_q^{(2)} - \epsilon_q^{(2)} = \langle \delta \tilde{\chi}_q^{(1)}, (H_0 - \epsilon_q) \delta \tilde{\chi}_q^{(1)} \rangle \geq 0 \quad (64)$$

for variations $\delta \tilde{\chi}_q^{(1)}$ for which

$$\langle \psi_k, \delta \tilde{\chi}_q^{(1)} \rangle = 0 \quad (k < q) \quad (65)$$

Whereas the exact solutions $\psi_q^{(1)}$ of the first order equations for different states q would yield wave functions orthogonal to the first order in λ , this is not generally the case for variational approximations. The conditions under which it is true are derived in Chapter V.

3. Variation Principle for Second Order Perturbation Equation

The variation principle of the Hylleraas type (53) for the second order perturbation function $\psi^{(2)}$ given by equation (50) is

$$\delta \tilde{\epsilon}^{(4)} \geq \epsilon^{(4)} \quad (66)$$

where

$$\begin{aligned} \tilde{\epsilon}^{(4)} = & \langle \tilde{\psi}^{(2)}, (H_0 - \epsilon) \tilde{\psi}^{(2)} \rangle + \langle \tilde{\psi}^{(2)}, V \psi^{(1)} \rangle + \langle \tilde{\psi}^{(1)}, V \tilde{\psi}^{(2)} \rangle \\ & - \epsilon^{(2)} [\langle \tilde{\psi}^{(2)}, \psi \rangle + \langle \psi, \tilde{\psi}^{(2)} \rangle], \end{aligned} \quad (67)$$

$\psi^{(1)}$ is the exact solution of (1) and $\tilde{\psi}^{(2)}$ is arbitrary. For cases in which (50) cannot be solved explicitly it is therefore possible to obtain a variational approximation for $\psi^{(2)}$ by making $\tilde{\epsilon}^{(4)}$ stationary with respect to parameters appearing in the trial function $\tilde{\psi}^{(2)}$. An obvious drawback to this procedure is that it requires an exact $\psi^{(1)}$, and if $\psi^{(1)}$ can be obtained explicitly then so in all likelihood can $\psi^{(2)}$ and the problem is not one requiring the variational approximation. The difficulty can be overcome in principle by adopting the approach of Chapter VI, which is to reformulate the second order stage in perturbation theory as a new first order problem.

C. Variational Principles for Lower Bounds to the Energy

In this section we shall discuss two other variation principles which, unlike the Ritz principle, give lower bounds to the energy of the ground state. The first of these, due to Temple (1928), applies to all orders of the energy, but has the disadvantage of requiring knowledge of the unperturbed energy of an excited state. The second, which has been discovered recently by Prager and Hirschfelder (1963), is specifically for the second order energy and does not require knowledge of an excited state.

1. Temple Principle

This principle (Temple, 1928 ; Hylleraas, 1961) is based on the inequality for the exact Hamiltonian

$$\langle (H - E_0) \tilde{\Psi}, (H - E_1) \tilde{\Psi} \rangle \geq 0 \quad (68)$$

where E_0 and E_1 are the two lowest eigenvalues of H (of a given symmetry) and $\tilde{\Psi}$ is any variation function. By means of the usual perturbation expansion we can deduce a second order result which may be combined with (53) to give

$$0 \leq \tilde{\epsilon}_0^{(2)} - \epsilon_0^{(2)} \leq \frac{\langle [(H_0 - \epsilon_0) \tilde{\psi}_0^{(1)} + V \psi_0]^2 \rangle}{\epsilon_1 - \epsilon_0} \quad (69)$$

where $\tilde{\epsilon}_0^{(2)}$ is given by (54). Prager and Hirschfelder (1963) have improved this upper bound further by optimizing with respect to the magnitude of $\tilde{\psi}_0^{(1)}$ which leads to

$$\epsilon_0^{(2)} \geq - \frac{\langle \psi_0, (V')^2 \psi_0 \rangle}{\epsilon_1 - \epsilon_0} + \frac{A_0^2}{B_0} \quad (70)$$

where

$$A_0 = \langle \tilde{\psi}_0^{(1)}, V' \psi_0 \rangle - \frac{\langle V' \psi_0, (H_0 - \epsilon_0) \tilde{\psi}_0^{(1)} \rangle}{\epsilon_1 - \epsilon_0} \quad (71)$$

$$B_0 = \frac{\langle (H_0 - \epsilon_0) \tilde{\psi}_0^{(1)}, (H_0 - \epsilon_0) \tilde{\psi}_0^{(1)} \rangle}{\epsilon_1 - \epsilon_0} - \langle \tilde{\psi}_0^{(1)}, (H_0 - \epsilon_0) \tilde{\psi}_0^{(1)} \rangle$$

The simpler inequality

$$\epsilon_0^{(2)} \geq - \frac{\langle \psi_0 (V')^2 \psi_0 \rangle}{\epsilon_1 - \epsilon_0} = - \sum_{k \geq 1} \frac{|V_{ok}|^2}{\epsilon_1 - \epsilon_0} \quad (72)$$

follows either by dropping the positive term A_0^2/B_0 or by setting $\tilde{\psi}_0^{(1)} = 0$. It can also be obtained from the exact perturbation expansion (2.20) by approximating the denominators $\epsilon_k - \epsilon_0$ by $\epsilon_1 - \epsilon_0$ and using the closure relation for the complete unperturbed set ψ_n ; it is originally due to Unsöld (1927). It is interesting to note³ that if this approximation is made in the sum (2.17) for the first order wave function we get the form (58),

$$\tilde{\psi}_0^{(1)} \approx - \frac{V'}{\epsilon_1 - \epsilon_0} \psi_0 \quad (73)$$

The Temple variation principle and the crude fixed lower bound (72) suffer from the disadvantage that they require a knowledge of the unperturbed energy ϵ_1 of the first excited state which is coupled to the ground state by the perturbation V . This may not be known.

2. Analogue of Thomson's Principle

The electrostatic analogy mentioned above has recently been fruitfully exploited by Prager and Hirschfelder (1963) to provide a variational lower bound for $\epsilon_0^{(2)}$ which does not require knowledge

of ϵ_1 . The principle in question is the analogue of Thomson's principle in electrostatics, and may be stated in the form

$$\epsilon_0^{(2)} \geq -\frac{1}{2} \sum_i \langle \psi_0, \tilde{G}_i \cdot \tilde{G}_i \psi_0 \rangle \quad (74)$$

where the trial vectors \tilde{G}_i satisfy the equation

$$\sum_i \nabla_i \cdot (\psi_0^2 \tilde{G}_i) = -2\psi_0 \nabla \cdot \psi_0 \quad (75)$$

and the boundary conditions

$$\psi_0^2 \tilde{G}_i = 0 \quad \text{on boundaries} \quad (76)$$

but are otherwise arbitrary. The equality holds if and only if

$\tilde{G}_i = -\nabla_i F$, when (74) reduces to (12). The proof is simple. Substitute $\tilde{G}_i = -\nabla_i F + \delta \tilde{G}_i$ into (74) to get

$$\tilde{\epsilon}_0^{(2)} = -\frac{1}{2} \sum_i \langle \psi_0, \tilde{G}_i \cdot \tilde{G}_i \psi_0 \rangle = \epsilon_0^{(2)} + \sum_i \langle \psi_0, \nabla_i F \cdot \delta \tilde{G}_i \psi_0 \rangle - \frac{1}{2} \sum_i \langle \psi_0, \delta \tilde{G}_i \cdot \delta \tilde{G}_i \psi_0 \rangle \quad (77)$$

But

$$\psi_0 \nabla_i F \cdot \delta \tilde{G}_i \psi_0 = \nabla_i \cdot (F \psi_0^2 \delta \tilde{G}_i) - F \nabla_i \cdot (\psi_0^2 \delta \tilde{G}_i) \quad (78)$$

The second term vanishes by (75) and the integral of the first term by (76). Hence

$$\tilde{\epsilon}_0^{(2)} = \epsilon_0^{(2)} - \frac{1}{2} \sum_i \langle \psi_0, \delta \tilde{G}_i \cdot \delta \tilde{G}_i \psi_0 \rangle$$

and the inequality is proved. In the three-dimensional case the difficulty of obtaining a vector solution of (75) can be overcome by putting

$$\tilde{\mathbf{E}} = \psi_0^{-2} [\nabla(\Phi + \phi) + \underline{\mathbf{D}}] \quad (79)$$

where Φ is a solution of the Poisson equation

$$\nabla^2 \Phi = 2\psi_0 (\nabla \cdot \mathbf{E}^{(1)}) \psi_0 \quad (80)$$

given by

$$\Phi(\underline{\mathbf{r}}) = -\frac{1}{2\pi} \int \frac{\psi_0'(\nabla' \cdot \mathbf{E}_0^{(1)}) \psi_0'}{|\underline{\mathbf{r}} - \underline{\mathbf{r}}'|} d\tau', \quad (81)$$

ϕ is a solution of Laplace's equation chosen so that the boundary conditions $\nabla(\Phi + \phi) = 0$ are satisfied, and $\underline{\mathbf{D}}$ is a vector such that $\text{div} \underline{\mathbf{D}} = 0$ and which vanishes at the boundary, but is otherwise arbitrary.

FOOTNOTES TO CHAPTER III

1. This is possible if the perturbation V is a real Hermitian operator, which is usually the case. The only important exceptions are magnetic perturbations.
2. Attention must be paid to the behavior of the right-hand side of Eq. (38) at the lower limit when $r \rightarrow 0$. For example, assuming $\psi(0) \neq 0$, $V'_{lm}\psi \simeq r^l$ and $K \simeq r^{-k}$ at the origin, we must have $k \leq 2+l$ in order that $M(r)$ exists. It will usually be possible to find a solution $K(r)$ of Eq. (36) which behaves correctly at the origin.
3. The authors are indebted to Mr. W. J. Meath for this observation.

IV. EXPECTATION VALUES AND DOUBLE PERTURBATION THEORY

In this chapter the perturbation theory of expectation values of operators other than the energy is discussed. The most important properties of particular states of a system are of two kinds, and can be written as the expectation values of two different kinds of operators.

First order properties, such as permanent dipole and quadrupole moments, diamagnetic susceptibilities, charge densities at the nucleus, are the straight expectation values of operators W , which are usually the sum of one-electron operators. For a system with Hamiltonian H in a state with energy and normalized eigenfunction Ψ , the expectation value of W is simply

$$\langle W \rangle = \langle \Psi, W \Psi \rangle \quad (1)$$

Such expectation values can be brought within the framework of perturbation theory by introducing a fictitious Hamiltonian

$$\mathcal{H} = H + \mu W$$

with eigenvalues $\mathcal{E}(\mu)$, so that $\mathcal{E}(0) = E$. Then since $W = \partial \mathcal{H} / \partial \mu$, by the Hellmann-Feynman theorem,

$$\begin{aligned} \langle W \rangle &= \langle \Psi, \partial \mathcal{H} / \partial \mu \Psi \rangle, \\ &= (\partial \mathcal{E} / \partial \mu)_{\mu=0} \end{aligned} \quad (2)$$

Thus W can be regarded as a first-order perturbation energy $E^{(1)}$ in the perturbation expansion of \mathcal{E} :

$$\mathcal{E} = E + \mu E^{(1)} + \mu^2 E^{(2)} + \dots \quad (3)$$

Second order properties include electric polarizabilities, paramagnetic susceptibilities, optical rotary constants, chemical shifts, nuclear spin-spin coupling constants, nuclear shielding constants, long range Van der Waals interaction constants. These are all essentially second order energies $E^{(2)}$ associated with the Hamiltonian \mathcal{H} (which in this case will usually have physical significance) and can be written in form

$$\langle Q \rangle = \frac{1}{2} (\partial^2 E / \partial \mu^2)_{\mu=0} = \langle \chi, (W - \langle W \rangle) \Psi \rangle \quad (4)$$

where χ is the solution of the equation

$$(H - E) \chi + (W - \langle W \rangle) \Psi = 0 \quad (5)$$

Second order properties can be written formally as expectation values of operators involving the inverse of $(H - E)$; thus Q can be regarded as the operator

$$Q = - (W - \langle W \rangle) (H - E)^{-1} (W - \langle W \rangle) \quad (6)$$

Since the exact eigenfunctions Ψ of many-electron atoms and molecules are unknown, the calculation of first and second order properties starts with approximate wave functions ψ . The approximate expectation values calculated in this way compare unfavorably in accuracy with the energy, because whereas knowledge of an approximate wave function ψ is sufficient to calculate the energy through first order in the error, this is not true in general for any other operator. The corrections to such approximate expectation values are therefore a matter of importance. The Hamiltonian H can be written in the form

$$H = H_0 + \lambda V$$

where H_0 is the approximate (unperturbed) Hamiltonian corresponding to the approximate wave function $\psi (= \psi^{(0)})$ and λV is the remainder which is responsible for the corrections. Since expectation values can be regarded as perturbation energies associated with an operator W , the appropriate tool for calculating the corrections in a systematic manner is a double perturbation theory based on the Hamiltonian

$$\hat{H} = H_0 + \lambda V + \mu W \quad (7)$$

This approach has been used extensively by Dalgarno and his collaborators (Dalgarno and Lewis, 1955 ; Dalgarno and Stewart, 1958b) and is the subject of the present chapter.

Attention is naturally focussed on the leading corrections of first order in the actual or fictitious perturbation parameter λ . At first sight even these seem impossibly difficult to obtain explicitly because they involve perturbed wave functions such as $\psi^{(1)}$, given by Eq. (3.1) for the case in which the perturbation involves the two-electron repulsion terms. Fortunately it is not in fact necessary to find such wave functions in order to evaluate the first order corrections to expectation values, as the corrections can be expressed entirely in terms of integrals involving only the accessible solutions of first order perturbation equations with one-electron perturbation terms. The theorem permitting this alternative form has been exploited very thoroughly for atoms by Dalgarno and his collaborators (1955, 1958b, 1960), and together with the integrability of the one-electron

first-order equations has allowed them to evaluate the leading corrections to the expectation values for a variety of operators. For this reason we refer to the relation as Dalgarno's Interchange Theorem; it is derived in Section B of this Chapter.

Schwartz (1959) has developed what appears to be a different approach in that he deliberately avoids the use of a Hamiltonian H_0 , corresponding to the approximate wave function ψ_0 and bases his method on a variation principle (see Chapter V). However, since he too seeks to calculate $\langle W \rangle$ and $\langle Q \rangle$ correct to the first order in $\Psi - \psi$, the results are equivalent to those of double perturbation theory, as we shall show. It should be emphasized that, as stated in the Introduction, a Hamiltonian H_0 can always be constructed and used in the formal development of perturbation theory. An interesting feature revealed by Schwartz's approach is that H_0 may then be eliminated from the final formulae if desired.

A. Double Perturbation Theory

The Rayleigh-Schrodinger perturbation formulae of Chapter II are easy to generalize to double perturbation problems where the Hamiltonian \mathcal{H} is given by Eq. (7). The basic assumption is that the wave function Φ and energy \mathcal{E} for the perturbed state can be expanded in a double power series in λ and μ ,

$$\Phi = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \lambda^n \mu^m \psi^{(n,m)}, \quad (8)$$

$$\mathcal{E} = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \lambda^n \mu^m \epsilon^{(n,m)} \quad (9)$$

The resulting perturbation equations for $\psi^{(n,0)}$ and $\epsilon^{(n,0)}$ are the same as Eqs. (2.3) through (2.10) with $\psi^{(n,0)}$ replacing $\psi^{(n)}$ and $\epsilon^{(n,0)}$ replacing $\epsilon^{(n)}$. Similarly, the relations for $\psi^{(0,n)}$ and $\epsilon^{(0,n)}$ are the same as Eqs. (2.3) through (2.10) with $\psi^{(0,n)}$ replacing $\psi^{(n)}$, $\epsilon^{(0,n)}$ replacing $\epsilon^{(n)}$, W replacing V and $W' = W - \epsilon^{(0,1)}$ replacing V' . It is only the mixed perturbation equations which are essentially new, the general member being

$$H_0 \psi^{(n,m)} + V \psi^{(n-1,m)} + W \psi^{(n,m-1)} = \sum_{j=0}^n \sum_{k=0}^m \epsilon^{(j,k)} \psi^{(n-j,m-k)} \quad (10)$$

In order that Φ be normalized to unity for all values of λ and μ , the perturbed wave functions $\psi^{(j,k)}$ must satisfy the conditions, analogous to Eq. (2.5),

$$\sum_{j=0}^n \sum_{k=0}^m \langle \psi^{(j,k)}, \psi^{(n-j,m-k)} \rangle = \delta_{n0} \delta_{m0}, \quad n, m = 0, 1, 2, \dots \quad (11)$$

A general expression for the mixed order $\epsilon^{(n,m)}$ can be obtained by multiplying Eq. (2.22) by ψ^* and integrating over all space,

$$\begin{aligned} \epsilon^{(n,m)} &= \langle \psi, W \psi^{(n-1,m)} \rangle + \langle \psi, W \psi^{(n,m-1)} \rangle \\ &- \sum_{j=0}^n \sum_{k=0}^m \epsilon^{(j,k)} [1 - \delta_{j0} \delta_{k0} - \delta_{jn} \delta_{km}] \langle \psi, \psi^{(n-j,m-k)} \rangle \end{aligned} \quad (12)$$

These double perturbation energies can be expressed in a variety of forms, some of which are discussed in Section B.

1. Expectation Values

The first and second order properties $\langle W \rangle$ and $\langle Q \rangle$ associated with an operator W can be expanded in powers of the

perturbation parameter λ , and the coefficients expressed in terms of the perturbation energies $\epsilon^{(n,m)}$. The general connection can be made by comparing Eqs. (3) and (9), which leads to the expansions

$$E^{(m)} = \sum_{n=0}^{\infty} \lambda^n \epsilon^{(n,m)}, \quad m = 0, 1, \dots \quad (13)$$

The corresponding perturbed wave functions are given by

$$\Psi^{(m)} = \sum_{n=0}^{\infty} \lambda^n \psi^{(n,m)}, \quad m = 0, 1, \dots \quad (14)$$

(a) First Order Properties. The first order property $\langle W \rangle$ defined by Eq. (1) can be expanded in powers of λ by means of Eq. (14):

$$\langle W \rangle = \sum_{n=0}^{\infty} \lambda^n \sum_{k=0}^n \langle \psi^{(k,0)}, W \psi^{(n-k,0)} \rangle, \quad (15a)$$

$$= \langle \psi, W \psi \rangle + \lambda [\langle \psi^{(1,0)}, W \psi \rangle + \langle \psi, W \psi^{(1,0)} \rangle] + \dots \quad (15b)$$

In this form the first order correction involves $\psi^{(1,0)}$, which is the solution of the equation

$$(H_0 - \epsilon) \psi^{(1,0)} + V' \psi = 0 \quad (16)$$

that has been made orthogonal to ψ . Equation (15) is useless for explicit calculation since Eq. (16) cannot be solved in closed form if V contains two-particle electron repulsion terms. Note that since $W = E^{(1)}$, the corrections in Eq. (15) can also be written in the form

$$\langle W \rangle = \sum_{n=0}^{\infty} \lambda^n \epsilon^{(n,1)} \quad (17)$$

(b) Second Order Properties. The second order property $\langle Q \rangle$, defined by Eq. (2) can be expanded in powers of λ by means of Eq. (14):

$$\langle Q \rangle = \sum_{n=0}^{\infty} \lambda^n \sum_{k=0}^n \langle \psi^{(k,1)}, W \psi^{(n-k,0)} \rangle, \quad (18a)$$

$$= \langle \psi^{(0,1)}, W \psi \rangle + \lambda [\langle \psi^{(1,1)}, W \psi \rangle + \langle \psi^{(1,0)}, W \psi^{(1,0)} \rangle] + \dots \quad (18b)$$

where $\psi^{(0,1)}$ is the solution of

$$(H_0 - \epsilon) \psi^{(0,1)} + W \psi = 0 \quad (19a)$$

and $\psi^{(1,1)}$ is the solution of Eq. (10) with $n = m = 1$,

$$(H_0 - \epsilon) \psi^{(1,1)} + V \psi^{(0,1)} + W \psi^{(1,0)} = 0 \quad (20)$$

The initial approximation for $\langle Q \rangle$, given by the first term of Eq. (18b), has the same form as a second order energy, and requires for its evaluation the solution of Eq. (19a). As discussed in Chapter IIIA, an equation of this type can frequently be solved explicitly when W is a one-particle operator by putting $\psi^{(0,1)} = F \psi$, so that it assumes the form of Eq. (3.3), namely

$$[H_0, F] \psi + W \psi = 0 \quad (19b)$$

If $H - H_0 = \lambda V$ is a function of coordinates only we can go one step further and eliminate H_0 to get Schwartz's (1959) form¹

$$[H, F] \psi + W \psi = 0 \quad (19c)$$

$$\langle Q \rangle = \sum_{n=0}^{\infty} \lambda^n \langle \psi^{(n,1)}, W \psi \rangle$$

As far as the correction terms are concerned, however, Eq. (20) is no easier to solve than Eq. (16). Thus the form (18) is of little practical interest for calculating the correction terms. Note that since $\langle Q \rangle = E^{(2)}$ the corrections may also be written in the form

$$\langle Q \rangle = \sum_{n=0}^{\infty} \lambda^n \epsilon^{(n,2)} \quad (21)$$

B. Dalgarno's Interchange Theorem

The general form of this theorem (Dalgarno and Stewart, 1958b) allows one to express the mixed perturbation energies $\epsilon^{(1,m)}$ entirely in terms of the W-perturbed functions $\psi^{(0,m)}$. To establish the connection it is convenient to introduce the Hamiltonian

$$H_* = H_0 + \mu W \quad (22)$$

with eigenfunctions χ and eigenvalues E_* . The perturbation expansions of χ (normalized) and E_* are given by Eqs. (8) and (9) as

$$\chi = \sum_{m=0}^{\infty} \mu^m \psi^{(0,m)} \quad (23)$$

$$E_* = \sum_{m=0}^{\infty} \mu^m \epsilon^{(0,m)} \quad (24)$$

Let us now regard \mathcal{H} as the single perturbation Hamiltonian

$$\mathcal{H} = H_* + \lambda V \quad (25)$$

The first order energy is simply

$$\begin{aligned}
 E_*^{(1)} &= \langle \chi, V \chi \rangle, \\
 &= \sum_{m=0}^{\infty} \mu^m \sum_{k=0}^m \langle \psi^{(0,m-k)}, V \psi^{(0,k)} \rangle.
 \end{aligned}
 \tag{26}$$

But by the definition of $E_*^{(1)}$ it follows from (9) that

$$E_*^{(1)} = \sum_{m=0}^{\infty} \mu^m \epsilon^{(1,m)}, \tag{27}$$

and therefore by comparing Eqs. (26) and (27),

$$\epsilon^{(1,m)} = \sum_{k=0}^m \langle \psi^{(0,m-k)}, V \psi^{(0,k)} \rangle. \tag{28}$$

This proves that the integrals in Eqs. (15) and (18) involving $W \psi^{(1,m)}$ can be interchanged for others involving $V \psi^{(0,m+1)}$, which is the content of the Interchange Theorem.

(a) First Order Properties. By using Eq. (28) for $\epsilon^{(1,0)}$ the expectation value of W through the first order can be written in the form (Dalgarno and Stewart, 1956b)

$$\langle W \rangle = \langle \psi, W \psi \rangle + \lambda [\langle \psi^{(0,1)}, V \psi \rangle + \langle \psi, V \psi^{(0,1)} \rangle] + \dots, \tag{29a}$$

where $\psi^{(0,1)}$ is the solution of Eq. (19). But putting $\psi^{(0,1)} = F \psi$ and $\lambda V = H - \bar{H}_0$ Eq. (29a) can be written in the form derived by Schwartz (1959)

$$\langle W \rangle \approx \langle \psi, [W + F(H - \bar{H}) + (H - \bar{H})F] \psi \rangle, \tag{29b}$$

where F is given by Eq. (19c) and, following Schwartz, $\bar{H} = \langle \psi, H \psi \rangle$ has been introduced to free F from the usual orthogonality restriction

$\langle \psi, F\psi \rangle = 0$. Since Eq. (19) can usually be solved explicitly, and the integrations in Eqs. (29) are straightforward, $\langle W \rangle$ can now be calculated exactly to the same order of accuracy as the energy.

(b) Second Order Properties. In a similar fashion the use of Eq. (28) for $\epsilon^{(2,1)}$ allows the second order property $\langle Q \rangle$ to be written in the form

$$\langle Q \rangle = \langle \psi^{(0,1)}, W\psi \rangle + \lambda [\langle \psi^{(0,2)}, W\psi \rangle + \langle \psi, W\psi^{(0,2)} \rangle + \langle \psi^{(0,1)}, V\psi^{(0,1)} \rangle] + \dots, \quad (30a)$$

where the function $\psi^{(0,2)}$ is the solution of the second order single perturbation equation

$$(H_0 - \epsilon)\psi^{(0,2)} + W'\psi^{(0,1)} - \epsilon^{(0,2)}\psi = 0. \quad (31a)$$

Schwartz's (1959) form of Eq. (30a) can be obtained in a similar fashion to Eq. (29b) by putting $\psi^{(0,2)} = G\psi$, etc., and making the additional assumption that V is a function of coordinates only:

$$\langle Q \rangle \approx \langle \psi, [W'F + G(H - \bar{H}) + (H - \bar{H})G + |F|^2(H - \bar{H})]\psi \rangle. \quad (30b)$$

Equation (31a) can be written in terms of G and F , and H_0 can be eliminated in favor of H , to give the equation quoted by Schwartz:

$$[H, G]\psi + W'F\psi - \epsilon^{(0,2)}\psi = 0. \quad (31b)$$

Since this equation can also frequently be solved explicitly by the techniques of Chapter III, the first order corrections to $\langle Q \rangle$ can be calculated exactly in a large number of cases. This result has been very widely applied to the calculation of the second order

properties of atoms by many authors (see Dalgarno, 1962).

C. Solution of Perturbation Equations

The methods of solving Eqs. (19) and (31) when W is a one-electron operator have been thoroughly discussed in Chapter IIIA. These equations correspond to Eqs. (3.1) and (3.50), the only differences being the notational ones: $W' \rightarrow V'$, $\psi^{(0,1)} \rightarrow \psi^{(1)}$, $\psi^{(0,2)} \rightarrow \psi^{(2)}$, $\epsilon^{(0,2)} \rightarrow \epsilon^{(2)}$. When explicit solutions for $\psi^{(0,1)}$ and $\psi^{(0,2)}$ cannot be found, various approximate variational techniques can be used. The straightforward method, discussed in Chapter IIIB, is to use the Hylleraas variation principles, Eqs (3.54) and (3.67). Modifications have also been used (Karplus and Kolker, 1963a) and are discussed briefly in Chapter XI.

The only cases in which explicit formulae can be given for the corrections are when the equations are one-dimensional or separable. In the one-dimensional case discussed in Chapter III, Section A2, the first order correction $\epsilon^{(1,1)}$ can be expressed in a form similar to Eq. (3.26) which applies to ground or excited states, namely

$$\epsilon^{(1,1)} = - \int_a^b \frac{M(z)N(z)}{\psi^2(z)} dz, \quad (32)$$

where

$$M(z) = 2 \int_a^z \psi V' \psi dz, \quad (33)$$

$$N(z) = 2 \int_a^z \psi W' \psi dz. \quad (34)$$

A similar type of formula can be derived when H and W are separable in the same coordinate system (see Appendix D).

D. Expansion in Perturbed and Unperturbed Eigenfunctions

In this section the expression of the first and second order properties $\langle W \rangle$ and $\langle Q \rangle$ in terms of the perturbed eigenfunctions Ψ_k of H (single perturbation) and in terms of the unperturbed eigenfunctions ψ_k of H_0 (double perturbation) are presented.

We begin by considering the difference between first and second order properties. Whereas the expectation value of W for a particular state q of a system with Hamiltonian H requires only a knowledge of Ψ_q , namely

$$\langle W \rangle_q = \langle \Psi_q, W \Psi_q \rangle, \quad (35)$$

the formal expression for the second order property $\langle Q \rangle$ in terms of the solutions of H requires the entire spectrum of eigenfunctions and eigenvalues. This can be seen by expanding χ , given by Eq. (5), in terms of the eigenfunctions Ψ_k and substituting into Eq. (4). The resulting well known expression has exactly the same form as the second order energy expansion Eq. (2.20) and is

$$\langle Q \rangle_q = \sum_k' \frac{W_{qk} W_{kq}}{E_q - E_k}, \quad (36)$$

where

$$W_{qk} = \langle \Psi_q, W \Psi_k \rangle, \quad (37)$$

the sum is to be taken to include integration over the continuum, and the prime indicates that states with energy E_q are to be omitted.

It is instructive to compare these formally exact expressions with the approximate expressions obtained when $H = H_0 + \lambda V$ and we

regard only the spectrum of the unperturbed Hamiltonian H_0 as known. By expanding the perturbation functions $\psi^{(0,1)}$ and $\psi^{(0,2)}$ of Eqs. (19) and (31) in terms of the unperturbed functions ψ_k and using Eqs. (29) and (30) we obtain

$$\langle W \rangle_q = W_{qq}^{(0)} + 2\lambda \sum_k \frac{W_{qk}^{(0)} V_{kq}^{(0)}}{\epsilon_q - \epsilon_k} + \dots, \quad (38)$$

$$\langle Q \rangle_q = 2 \sum_k \frac{W_{qk}^{(0)} W_{kq}^{(0)}}{\epsilon_q - \epsilon_k} + 2\lambda \sum_k \sum_j \frac{(V_{qk}^{(0)} W_{kj}^{(0)} W_{jq}^{(0)} + W_{qk}^{(0)} V_{kj}^{(0)} W_{jq}^{(0)} + W_{qk}^{(0)} W_{kj}^{(0)} V_{jq}^{(0)})}{(\epsilon_q - \epsilon_k)(\epsilon_q - \epsilon_j)}, \quad (39)$$

where

$$W_{qq}^{(0)} = \langle \psi_q, W \psi_q \rangle,$$

$$W_{kj}^{(0)} = \langle \psi_k, W \psi_j \rangle - \delta_{kj} \langle \psi_q, W \psi_q \rangle, \quad (40)$$

$$V_{kj}^{(0)} = \langle \psi_k, V \psi_j \rangle - \delta_{kj} \langle \psi_q, V \psi_q \rangle \quad (\text{all } k, j \text{ except } k = j = q).$$

The first order corrections due to the perturbation V are effected through the matrix elements $V_{kj}^{(0)}$ in the terms of order λ .

There are two important and related second order properties of electronic systems which we shall consider in more detail to exemplify these formulae. They are the polarizability α of an atom or molecule and the Van der Waals constant C for the long range interaction of two atoms or molecules.

The operator W for the polarizability tensor is the electric dipole moment vector

$$\mu = \sum_s e_s r_s \quad (41)$$

where e_s and r_s are the charge (see footnote 2, Chapter X) and position of particle s , and the summation is over all particles. The mean polarizability for a system in state q , given by Eq. (36), can be written (cf. Chapter X, where ϵ is used in the sense of the E in the present chapter)

$$\alpha_q = \sum_k \frac{f_{qk}}{(E_q - E_k)^2} \quad (42)$$

where f_{qk} is the mean oscillator strength defined by

$$f_{qk} = \frac{2}{3} (E_k - E_q) |\langle \Psi_k, \mu \Psi_q \rangle|^2 \quad (43)$$

If the atom or molecule is isotropic then

$$f_{qk} = 2(E_k - E_q) |(\mu_x)_{qk}|^2 \quad \text{when} \quad (\mu_x)_{qk} = \langle \Psi_q, \mu_x \Psi_k \rangle \quad (44)$$

and μ_x is one of the components of μ .

The zeroeth order approximation $\alpha_q^{(0)}$ to the polarizability of a system based on an approximate wave function $\psi_q^{(0)}$, such as a Hartree-Fock, which takes inadequate account of electron repulsion will be given by Eq. (38), namely

$$\alpha_q^{(0)} = \sum_k' \frac{f_{qk}^{(0)}}{(E_q - E_k)^2} \quad (45)$$

The leading correction for the perturbation V , which involves the two-particle electron repulsion terms, is given by

$$\alpha_q^{(1)} = 2 \sum_k \sum_j \frac{(\mu_{qk}^{(0)} v_{kj}^{(0)} \mu_{jq}^{(0)} + 2 \mu_{qk}^{(0)} \mu_{kj}^{(0)} v_{jq}^{(0)})}{(\epsilon_q - \epsilon_k)(\epsilon_q - \epsilon_j)} \quad (46)$$

where $v_{kj}^{(0)}$ is defined by Eq. (40) and

$$\mu_{kj}^{(0)} = \langle \psi_k, \mu_x \psi_j \rangle - \delta_{kj} \langle \psi_q, \mu_x \psi_q \rangle \quad (47)$$

The treatment of the Van der Waals constant C is similar (see Chapter X, Section C3). Consider the long range interaction of two identical atoms A and B in ground S states at distance R apart. The leading terms in the multipole expansion of the interaction potential is the two-particle perturbation

$$W = \frac{1}{R^3} (\mu_a^a \cdot \mu_b^b - 3 \mu_x^a \mu_x^b) \quad (48)$$

where μ_a^a and μ_b^b are the electric dipole operators for A and B and μ_x^a, μ_x^b are the components along the molecular axis. C is defined as the coefficient of R^{-6} in the expansion of the interaction energy in powers of $1/R$. The exact expression for C in terms of the atomic eigenfunctions Ψ_k is therefore

$$C = 3 \sum_k \sum_j \frac{f_{qk} f_{qj}}{(\epsilon_k - \epsilon_q)(\epsilon_j - \epsilon_q)(\epsilon_k + \epsilon_j - 2\epsilon_q)} \quad (49)$$

However if, as is the case for all atoms other than hydrogen, approximate atomic wave functions $\psi_k^{(0)}$ are used then C has correction terms of the same kind as (40) and involving the same matrix elements.

FOOTNOTES TO CHAPTER IV

1. This is similar to, but different from, Eq. (3.4), since here we are dealing with double perturbation theory, and H involves V and not W .

V. GENERAL CONNECTIONS BETWEEN PERTURBATION THEORY

AND THE VARIATIONAL PRINCIPLE

There are many connections between perturbation theory and the variational principle, some of which have been discussed in the preceding chapters. In this chapter we will explore certain of these connections in more detail.

First we want to emphasize that the relationship between perturbation theory and the variational principle is two-way. Consider some quantity A for which we wish a perturbation expression, $A = \sum_n \nu^n A^{(n)}$, where ν is the order parameter. Then, on the one hand, given a variational principle for A one can derive variational principles for the individual $A^{(n)}$. This possibility is obviously of great practical importance. We have already discussed the case $A = E$, in Chapter III and will discuss it further below. We would note however, that one can do similar things if A is a scattering phase shift or a scattering amplitude.

On the other hand the perturbation formula for $A^{(0)} + \nu A^{(1)}$ in fact provides a variational principle for A because the mark of the latter is precisely that it yields errors of second order.¹ We have already had one example of this sort in Chapter IV in discussing the connection between the double-perturbation and Schwartz approaches to the calculation of expectation values. Another, more familiar example² is the following: From first order perturbation theory we know that $\langle \psi^{(0)}, H \psi^{(0)} \rangle / \langle \psi^{(0)}, \psi^{(0)} \rangle = E + O(\nu^2)$ whence we infer that $\langle \tilde{\psi}, H \tilde{\psi} \rangle / \langle \tilde{\psi}, \tilde{\psi} \rangle$ is stationary for first order variations of $\tilde{\psi}$

about the true wave function, which is of course correct.³ As usual, we use tilde to represent trial functions.

A. Perturbation Theory of the Variational Principle

Let us now consider the case $A = E$ in some detail where, for notational simplicity, we continue to consider only a single perturbation which may be thought of either as λV or μH . Thus, we consider the Hamiltonian

$$\bar{H} = H_0 + \lambda H_1$$

and the associated stationary expression J where

$$J \equiv \langle \tilde{\psi} | (\bar{H} - \tilde{E}) | \tilde{\psi} \rangle \quad (1)$$

For fixed \tilde{E} one determines the optimal $\tilde{\psi}$ from $\delta J = 0$ and then one determines \tilde{E} from $J = 0$. We will symbolize this process by $\delta J = 0, J = 0$. To derive variational expressions for individual expansion coefficients we simply insert

$$J = \sum \lambda^n J^{(n)}, \quad \tilde{\psi} = \sum \lambda^n \tilde{\psi}^{(n)}, \quad \tilde{E} = \sum \lambda^n \tilde{E}^{(n)}$$

into Eq. (1) and equate terms of like power of λ to obtain the following sequence of stationary expressions.⁴

$$J^{(0)} \equiv \langle \tilde{\psi}^{(0)} | (H_0 - \tilde{E}^{(0)}) | \tilde{\psi}^{(0)} \rangle \quad (2)$$

$$J^{(1)} \equiv \langle \tilde{\psi}^{(0)} | (H_1 - \tilde{E}^{(1)}) | \tilde{\psi}^{(0)} \rangle + \langle \tilde{\psi}^{(1)} | (H_0 - \tilde{E}^{(0)}) | \tilde{\psi}^{(0)} \rangle + \langle \tilde{\psi}^{(1)} | (H_1 - \tilde{E}^{(1)}) | \tilde{\psi}^{(1)} \rangle \quad (3)$$

$$\begin{aligned}
 J^{(2)} \equiv & \langle \tilde{\psi}^{(0)}, (-\tilde{E}^{(2)}) \tilde{\psi}^{(0)} \rangle + \langle \tilde{\psi}^{(0)}, (\mathcal{H}_1 - \tilde{E}^{(1)}) \tilde{\psi}^{(1)} \rangle + \langle \tilde{\psi}^{(1)}, (\mathcal{H}_1 - \tilde{E}^{(1)}) \tilde{\psi}^{(0)} \rangle \\
 & + \langle \tilde{\psi}^{(1)}, (\mathcal{H}_0 - \tilde{E}^{(0)}) \tilde{\psi}^{(1)} \rangle + \langle \tilde{\psi}^{(0)}, (\mathcal{H}_0 - \tilde{E}^{(0)}) \tilde{\psi}^{(2)} \rangle + \langle \tilde{\psi}^{(2)}, (\mathcal{H}_0 - \tilde{E}^{(0)}) \tilde{\psi}^{(0)} \rangle
 \end{aligned}
 \tag{4}$$

$$J^{(3)} \equiv \text{etc.}$$

1. First Order Variation Principle

We will now examine the content of each of these stationary expressions in turn. Equation (2) is of course the variational expression for the zero order problem and requires no further comment.

(a) Delves' Variation Principle. Equation (3) is a variational principle for $E^{(1)}$. $J^{(1)}$ is stationary for variations of $\tilde{\psi}^{(1)}$ and $\tilde{\psi}^{(0)}$ about their correct values. Thus $\delta J^{(1)} = 0$, $J^{(1)} = 0$, $J^{(0)} = 0$ can provide approximate $E^{(1)}$, $\psi^{(1)}$ and $\psi^{(0)}$, the latter in general different from the approximate $\psi^{(0)}$ furnished by $\delta J^{(0)} = 0$, $J^{(0)} = 0$.

Until recently this variation principle has not attracted any attention. The reasons are simple:

(i) In the calculation of the energy levels of an isolated system by perturbation theory ($\mathcal{H}_0 = H_0$, $\mathcal{V}\mathcal{H}_1 = \lambda V$) one naturally chooses \mathcal{H}_0 so that $\psi^{(0)}$ and hence $E^{(1)}$ are known exactly and one has no need for $J^{(1)}$.

(ii) In the calculation of the effects of external fields ($\mathcal{H}_0 = H$, $\mathcal{V}\mathcal{H}_1 = \mu W$), very often $E^{(1)} = 0$ by symmetry whence, once again one has no need for $J^{(1)}$.

(iii) Given, as is usually the case, a definite but approximate

$\psi^{(0)}$, i.e. a $\tilde{\psi}^{(0)}$, then $J^{(1)}$ is likely to be of little help in determining an optimal $\psi^{(1)}$ because unrestricted variation of $\tilde{\psi}^{(1)}$ clearly yields an equation for $\tilde{\psi}^{(0)}$ not $\tilde{\psi}^{(1)}$.

However, as has been pointed out recently by Delves,⁵ if we do not know $\psi^{(0)}$ exactly and if one is interested in $\langle \psi^{(0)}, H_1 \psi^{(0)} \rangle$ for some H_1 , then this variational principle offers the possibility of an improvement over $\langle \tilde{\psi}^{(0)}, H_1 \tilde{\psi}^{(0)} \rangle$. More specifically the resultant $\tilde{E}^{(1)}$ will be correct to terms of order $\tilde{\psi}^{(0)} - \psi^{(0)}$ and $\tilde{\psi}^{(1)} - \psi^{(1)}$. However, this variational principle has at least two drawbacks:

(a) Even for the ground state it is not a minimal principle, only a stationary principle.⁶ Thus what we have termed an improvement may in fact not be one at all, i.e. may not be numerically more accurate, and in particular one does not necessarily improve things by using more elaborate trial functions.

(b) Related to this is the following: we are allowed to vary $\tilde{\psi}^{(0)}$ and $\tilde{\psi}^{(1)}$ independently and this can lead to peculiar results. In particular we may fix $\tilde{\psi}^{(0)}$ (determined say from $\delta J^{(0)} = 0$, $J^{(0)} = 0$) and vary only the scale of $\tilde{\psi}^{(1)}$ (i.e. we write $\tilde{\psi}^{(1)} = S\phi$ and vary the parameter S). One now readily finds that the optimal $\tilde{E}^{(1)}$ is simply $\langle \tilde{\psi}^{(0)}, H_1 \tilde{\psi}^{(0)} \rangle$, i.e. no improvement.

(b) Schwartz' Method. One may regard the Schwartz method (discussed in Chapter IV and shown there to be equivalent to the Double Perturbation Method) to be a special case of the Delves' variation principle in which, for fixed $\tilde{\psi}^{(0)}$ one restricts the freedom of $\tilde{\psi}^{(1)}$ by requiring that $\tilde{\psi}^{(0)} - \psi^{(0)}$ and $\tilde{\psi}^{(1)} - \psi^{(1)}$ be of the same order, so that the error in $\tilde{E}^{(1)}$ is of order $(\tilde{\psi}^{(0)} - \psi^{(0)})^2$.

This is done in the following way: writing $\psi^{(1)} = F\psi^{(0)}$ we have the Dalgarno-Lewis equation for F :

$$[\mathcal{H}_0, F]\psi^{(0)} = -(\mathcal{H}_1 - \langle \psi^{(0)}, \mathcal{H}_1 \psi^{(0)} \rangle) \psi^{(0)}$$

If now we write $\tilde{\psi}^{(1)} = \tilde{F}\tilde{\psi}^{(0)}$ it is clear that the error in $\tilde{\psi}^{(1)}$ will be of the same order as that in $\tilde{\psi}^{(0)}$ if \tilde{F} satisfies

$$[\mathcal{H}_0, \tilde{F}]\tilde{\psi}^{(0)} = -(\mathcal{H}_1 - \langle \tilde{\psi}^{(0)}, \mathcal{H}_1 \tilde{\psi}^{(0)} \rangle) \tilde{\psi}^{(0)} \quad (5)$$

which is the Schwartz condition. As Schwartz notes, one can also derive this equation from a variation principle (see below), which then can be used, if necessary, to effect an approximate solution.

2. Second Order Variational Principle

Equation (4) is a variational principle for $E^{(2)}$. In this general form it has not been used though it could have applicability to external field problems where usually one does not know $\psi^{(0)}$ exactly. However, following Schwartz, it is probably better to restrict $\tilde{\psi}^{(1)}$ and $\tilde{\psi}^{(2)}$ or following Dalgarno, to use double perturbation theory.

If $\psi^{(0)}$ is known exactly, i.e. $\tilde{\psi}^{(0)} = \psi^{(0)}$, $\tilde{E}^{(0)} = E^{(0)}$, $\tilde{E}^{(1)} = E^{(1)}$, then, so far as the numerical value of $J^{(2)}$ and its behavior under variations of $\tilde{\psi}^{(1)}$ are concerned we may drop the last

two terms whence $\delta J^{(2)} = 0$, $J^{(2)} = 0$ is identical to the Hylleraas minimal⁶ principle for $E^{(2)}$ and $\psi^{(1)}$. An essential point here is that, in contrast to the situation with $J^{(1)}$, free variation of $\tilde{\psi}^{(1)}$ does yield an equation for $\tilde{\psi}^{(1)}$. Similarly free variation of $\tilde{\psi}^{(n)}$ in $J^{(2n)}$ yields an equation for $\tilde{\psi}^{(n)}$ whence, if one knows the lower orders exactly, one can derive from $J^{(2n)}$ a Hylleraas-like variational principle⁷ which determines an optimal $\tilde{\psi}^{(n)}$ and an upper bound to $E^{(2n)}$.⁶

If on the other hand $\tilde{\psi}^{(0)}$ is only approximate then the expression $\delta \tilde{J}_H = 0$ with

$$\begin{aligned} \tilde{J}_H = & \langle \tilde{\psi}^{(0)}, (\mathcal{H}_1 - \tilde{E}^{(1)}) \tilde{\psi}^{(1)} \rangle + \langle \tilde{\psi}^{(1)}, (\mathcal{H}_1 - \tilde{E}^{(1)}) \tilde{\psi}^{(0)} \rangle \\ & + \langle \tilde{\psi}^{(1)}, (\mathcal{H}_0 - \tilde{E}^{(0)}) \tilde{\psi}^{(1)} \rangle \end{aligned}$$

with $\tilde{E}^{(0)} = \langle \tilde{\psi}^{(0)}, \mathcal{H}_0 \tilde{\psi}^{(0)} \rangle$ and $\tilde{E}^{(1)}$ approximated by $\langle \tilde{\psi}^{(0)}, \mathcal{H}_1 \tilde{\psi}^{(0)} \rangle$, can be used to determine an optimal $\tilde{\psi}^{(1)}$ for a given $\tilde{\psi}^{(0)}$. This $\tilde{\psi}^{(1)}$ can then be used in conjunction with $J^{(1)} = 0$ to yield a variational approximation to $E^{(1)}$.⁸ Indeed if we write $\tilde{\psi}^{(1)} = \tilde{R} \tilde{\psi}^{(0)}$ and make the approximation $\tilde{E}^{(0)} \tilde{\psi}^{(0)} \simeq \mathcal{H}_0 \tilde{\psi}^{(0)}$ then this becomes exactly the method of Schwartz. In this connection one should note that free variation of $\tilde{\psi}^{(1)}$ without the approximation $\tilde{E}^{(0)} \tilde{\psi}^{(0)} \simeq \mathcal{H}_0 \tilde{\psi}^{(0)}$ leads to an equation which one can presumably not solve exactly since, by hypothesis, one cannot solve the associated homogeneous equation. On the other hand one may hope to solve Schwartz' equation (5).

B. Special Theorems for Variational Wave Functions

1. Orthogonality and Related Theorems

We could now go on and discuss $J^{(3)}$ etc. but since no new questions of principle arise and the analysis becomes repetitive, we will terminate the detailed discussion at this point and turn to another question: Thus far we have been concerned with a single state. However, our variational procedures may well provide us with a number of stationary solutions which one would then be tempted to associate with various states. The question then arises--will these several solutions be orthogonal? In what follows we will, among other things, derive sufficient conditions for orthogonality. We believe that they are also necessary conditions if the orthogonality is to be enforced only by the variational principle, i.e. if the solutions in question are not a priori orthogonal by construction or because of some symmetry property.

Our basic tool is the observation that if $\tilde{\psi}_p$ and $\tilde{\psi}_q$ are optimal trial functions then from $\delta J = 0$ we have (varying $\tilde{\psi}_p$ but not $\tilde{\psi}_p^*$)

$$\tilde{E}_p \langle \tilde{\psi}_p, \delta \tilde{\psi}_p \rangle = \langle \tilde{\psi}_p, \tilde{H} \delta \tilde{\psi}_p \rangle \quad (6)$$

and (varying $\tilde{\psi}_q^*$ but not $\tilde{\psi}_q$)

$$\tilde{E}_q \langle \delta \tilde{\psi}_q, \tilde{\psi}_q \rangle = \langle \delta \tilde{\psi}_q, \tilde{H} \tilde{\psi}_q \rangle \quad (7)$$

From now on we will assume $\tilde{E}_p \neq \tilde{E}_q$. Then we have the following theorem:

(a) If $\delta\tilde{\psi}_p = \gamma\tilde{\psi}_q$ and $\delta\tilde{\psi}_q = \gamma\tilde{\psi}_p$ are possible variations* of the optimal trial functions then $\tilde{\psi}_p$ and $\tilde{\psi}_q$ are orthogonal.

Proof: The proof is immediate. Simply insert the variations into Eqs. (6) and (7) and subtract to find

$$(\tilde{E}_p - \tilde{E}_q) \langle \tilde{\psi}_p, \tilde{\psi}_q \rangle = 0$$

which proves the theorem.

From this follow as special cases, for example the orthogonality of the exact ψ_p and ψ_q and of different solutions of the Rayleigh-Ritz method. It is also of interest to make the following two remarks:

(b) Generalized Brillouin Theorem: If $\delta\tilde{\psi}_q = \gamma\psi'$ where $\langle \psi', \tilde{\psi}_q \rangle = 0$ is a possible variation of the optimal trial function then

$$\langle \psi', \tilde{H}\tilde{\psi}_q \rangle = 0$$

This is of course an immediate consequence of Eq. (7).

(c) Off-Diagonal Hypervirial Theorems for Variational Wave Functions:

If $\delta\tilde{\psi}_p = i\gamma W\tilde{\psi}_q$ and $\delta\tilde{\psi}_q = i\gamma W\tilde{\psi}_p$ where W is a Hermitian operator are allowed variations of the optimal trial functions then it follows by subtracting Eq. (6) and Eq. (7) that

$$(\tilde{E}_p - \tilde{E}_q) \langle \tilde{\psi}_p, W\tilde{\psi}_q \rangle = \langle \tilde{\psi}_p, [\tilde{H}, W]\tilde{\psi}_q \rangle$$

This generalizes an earlier result of Epstein and Hirschfelder (1961) which applies to the diagonal case $p = q$.

From this for example one can immediately infer, in agreement with other authors,⁹ that although all one-electron diagonal hypervirial theorems are satisfied in the Hartree-Fock approximation for closed

shells (see Chapter XI), off-diagonal ones are not.

Returning now to the problem of orthogonality, by expanding everything in powers of ψ , we can derive conditions that variational calculations based on $\delta J^{(0)} = \delta J^{(1)} = \delta J^{(2)} \dots = 0$ yield orthogonality to a certain order. This is of interest, for example, in connection with perturbation theory within the Hartree-Fock formalism (see Chapter XI) where in effect $\tilde{\psi}^{(0)}, \tilde{\psi}^{(1)}, \dots$ are the Hartree-Fock approximations to $\psi^{(0)}, \psi^{(1)}, \dots$.

Since the manipulations are straight forward we merely quote some of the results.

(d) If $\delta \tilde{\psi}_p^{(0)} = \eta \tilde{\psi}_q^{(0)}$ and $\delta \tilde{\psi}_q^{(0)} = \eta \tilde{\psi}_p^{(0)}$ are allowed variations then from $\delta J^{(0)} = 0$ we find $\langle \tilde{\psi}_p^{(0)}, \tilde{\psi}_q^{(0)} \rangle = 0$, i.e. we have orthogonality to zero order.

(e) If $\delta \tilde{\psi}_p^{(0)} = \eta \tilde{\psi}_q^{(0)}$ and $\delta \tilde{\psi}_q^{(0)} = \eta \tilde{\psi}_p^{(0)}$ are allowed variations, and if $\tilde{\psi}_p^{(0)} = \psi_p^{(0)}$ and $\tilde{\psi}_q^{(0)} = \psi_q^{(0)}$ then from $\delta J^{(1)} = 0$ we find we have orthogonality to first order, i.e. $\langle \tilde{\psi}_p^{(0)}, \tilde{\psi}_q^{(1)} \rangle + \langle \tilde{\psi}_p^{(1)}, \tilde{\psi}_q^{(0)} \rangle = 0$.

(f) If $\delta \tilde{\psi}_p^{(0)} = \eta \tilde{\psi}_q^{(0)}$, $\delta \tilde{\psi}_q^{(0)} = \eta \tilde{\psi}_p^{(0)}$, $\delta \tilde{\psi}_p^{(1)} = \eta \tilde{\psi}_q^{(1)}$ and $\delta \tilde{\psi}_q^{(1)} = \eta \tilde{\psi}_p^{(1)}$ are allowed variations and if $\tilde{\psi}_p^{(0)} = \psi_p^{(0)}$ and $\tilde{\psi}_q^{(0)} = \psi_q^{(0)}$ then from $\delta J^{(2)} = 0$ we find we have orthogonality to second order.

(g) etc.

As one application of these results, it follows directly in agreement with other authors,¹⁰ that in Hartree-Fock calculations of say the $(1s)^2$'s and $(1s2s)$'s states of He as an expansion in $1/2$,¹¹ one will have orthogonality to first order but not in higher orders.

These results, as noted are based on $\delta J^{(0)} = \delta J^{(1)} = \dots = 0$.

It is also clearly of interest to ask what happens if we use say the Hylleraas variational principle, i.e. $\delta J_p^H = 0$, $\delta J_q^H = 0$ where

$$J_p^H = \langle \psi_p^{(0)}, (\mathcal{H}_1 - E_p^{(1)}) \tilde{\psi}_p^{(1)} \rangle + \langle \tilde{\psi}_p^{(1)}, (\mathcal{H}_1 - E_p^{(1)}) \psi_p^{(0)} \rangle \\ + \langle \tilde{\psi}_p^{(1)}, (\mathcal{H}_0 - E_p^{(0)}) \tilde{\psi}_p^{(1)} \rangle$$

and a similar expression for J_q^H .

From $\delta J_p^H = \delta J_q^H = 0$ we then infer

$$0 = \langle \psi_p^{(0)}, (\mathcal{H}_1 - E_p^{(1)}) \delta \tilde{\psi}_p^{(1)} \rangle + \langle \tilde{\psi}_p^{(1)}, (\mathcal{H}_0 - E_p^{(0)}) \delta \tilde{\psi}_p^{(1)} \rangle$$

and

$$0 = \langle \delta \tilde{\psi}_q^{(1)}, (\mathcal{H}_1 - E_q^{(1)}) \psi_q^{(0)} \rangle + \langle \delta \tilde{\psi}_q^{(1)}, (\mathcal{H}_0 - E_q^{(0)}) \tilde{\psi}_q^{(1)} \rangle$$

By use of $\mathcal{H}_0 \psi_p^{(0)} = E_p^{(0)} \psi_p^{(0)}$, $\mathcal{H}_0 \tilde{\psi}_q^{(0)} = E_q^{(0)} \tilde{\psi}_q^{(0)}$ and $\langle \psi_p^{(0)}, \psi_q^{(0)} \rangle = 0$ we can then immediately derive, by subtracting the two expressions above, the following theorem:¹²

(h) If $\delta \tilde{\psi}_p^{(1)} = \gamma \psi_q^{(0)}$ and $\delta \tilde{\psi}_q^{(1)} = \eta \psi_p^{(0)}$ are possible variations of the optimal trial functions in the Hylleraas variational principle then we have orthogonality to first order.

In a similar way we can show:

(i) If $\delta \tilde{\psi}_p^{(1)} = i\eta W \psi_q^{(0)}$ and $\delta \tilde{\psi}_q^{(1)} = i\eta W \psi_p^{(0)}$ are allowed variations in the Hylleraas variation principle then the off diagonal hypervirial theorem for W will be satisfied to first order.

(j) If $\delta \tilde{\psi}_p^{(1)} = i\eta W \psi_p^{(0)}$ is an allowed variation in the Hylleraas variation principle then the hypervirial theorem for W will be satisfied to first order.

2. The Hellmann-Feynman Theorem

In this section we will discuss the Hellmann-Feynman theorem within the framework of the Hylleraas variational principle.¹³

Dropping the state subscript, we will show that if $\delta\tilde{\psi}^{(1)} = \gamma \frac{\partial\psi^{(0)}}{\partial\sigma}$ is an allowed variation where σ is the parameter in question then the Hellmann-Feynman theorem is satisfied to first order:

The Hellmann-Feynman theorem states that

$\frac{\partial E}{\partial\sigma} = \langle \Psi, \frac{\partial \mathcal{H}}{\partial\sigma} \Psi \rangle / \langle \Psi, \Psi \rangle$. In zero order we have then $\frac{\partial E^{(0)}}{\partial\sigma} = \langle \psi^{(0)}, \frac{\partial \mathcal{H}_0}{\partial\sigma} \psi^{(0)} \rangle$, which is certainly true. In first order then we find

$$\begin{aligned} \frac{\partial E^{(1)}}{\partial\sigma} &= \langle \psi^{(0)}, \frac{\partial \mathcal{H}_0}{\partial\sigma} \psi^{(0)} \rangle [-\langle \psi^{(0)}, \tilde{\psi}^{(1)} \rangle - \langle \tilde{\psi}^{(1)}, \psi^{(0)} \rangle] \\ &\quad + \langle \psi^{(0)}, \frac{\partial \mathcal{H}_0}{\partial\sigma} \tilde{\psi}^{(1)} \rangle + \langle \tilde{\psi}^{(1)}, \frac{\partial \mathcal{H}_0}{\partial\sigma} \psi^{(0)} \rangle + \langle \psi^{(0)}, \frac{\partial \mathcal{H}_1}{\partial\sigma} \psi^{(0)} \rangle \end{aligned} \quad (8)$$

On the other hand we have $E^{(1)} = \langle \psi^{(0)}, \mathcal{H}_1 \psi^{(0)} \rangle$ whence

$$\frac{\partial E^{(1)}}{\partial\sigma} = \langle \frac{\partial\psi^{(0)}}{\partial\sigma}, \mathcal{H}_1 \psi^{(0)} \rangle + \langle \psi^{(0)}, \mathcal{H}_1 \frac{\partial\psi^{(0)}}{\partial\sigma} \rangle + \langle \psi^{(0)}, \frac{\partial \mathcal{H}_1}{\partial\sigma} \psi^{(0)} \rangle \quad (9)$$

and the question is, are these two expressions equal?¹⁴

We now put $\delta J^H = 0$ with $\delta\tilde{\psi}^{(1)} = \gamma \frac{\partial\psi^{(0)}}{\partial\sigma}$ and find (varying both $\tilde{\psi}^{(1)}$ and $\tilde{\psi}^{(1)*}$)

$$\begin{aligned} 0 &= \langle \psi^{(0)}, (\mathcal{H}_1 - E^{(1)}) \frac{\partial\psi^{(0)}}{\partial\sigma} \rangle + \langle \frac{\partial\psi^{(0)}}{\partial\sigma}, (\mathcal{H}_1 - E^{(1)}) \psi^{(0)} \rangle \\ &\quad + \langle \frac{\partial\psi^{(0)}}{\partial\sigma}, (\mathcal{H}_0 - E^{(0)}) \tilde{\psi}^{(1)} \rangle + \langle \tilde{\psi}^{(1)}, (\mathcal{H}_0 - E^{(0)}) \frac{\partial\psi^{(0)}}{\partial\sigma} \rangle \end{aligned} \quad (10)$$

We now use $\frac{\partial}{\partial\sigma} \langle \psi^{(0)}, \psi^{(0)} \rangle = 0$,

$$(\mathcal{H}_0 - E^{(0)}) \frac{\partial \psi^{(0)}}{\partial \sigma} = \left(\frac{\partial E^{(0)}}{\partial \sigma} - \frac{\partial \mathcal{H}_0}{\partial \sigma} \right) \psi^{(0)}$$

which follows from differentiating $\mathcal{H}_0 \psi^{(0)} = E^{(0)} \psi^{(0)}$ with respect to σ , and the zero order Hellmann-Feynman Theorem to write Eq. (10) as

$$\begin{aligned} 0 &= \langle \psi^{(0)}, \mathcal{H}_1 \frac{\partial \psi^{(0)}}{\partial \sigma} \rangle + \langle \frac{\partial \psi^{(0)}}{\partial \sigma}, \mathcal{H}_1 \psi^{(0)} \rangle \\ &\quad - \langle \psi^{(0)}, \frac{\partial \mathcal{H}_0}{\partial \sigma} \tilde{\psi}^{(1)} \rangle - \langle \tilde{\psi}^{(1)}, \frac{\partial \mathcal{H}_0}{\partial \sigma} \psi^{(0)} \rangle \\ &\quad + \langle \psi^{(0)}, \frac{\partial \mathcal{H}_0}{\partial \sigma} \psi^{(0)} \rangle [\langle \psi^{(0)}, \tilde{\psi}^{(1)} \rangle + \langle \tilde{\psi}^{(1)}, \psi^{(0)} \rangle] \end{aligned}$$

which when inserted into Eq. (9) yields Eq. (8). Q.E.D.

FOOTNOTES TO CHAPTER V

1. R. E. B. Makinson and J. S. Turner, Proc. Phys. Soc. (London), 56 857 (1953).
2. For examples from scattering theory see footnote (1).
3. Further by using higher order perturbation theory one can derive "Super-Stationary" variation principles. See T. Kikuta, Prog. Theo. Phys. 12, 10 (1954); 14, 1453 (1955); L. Biedenharn, and J. M. Blatt, Phys. Rev. 93, 230 (1954) and Chapter VI.
4. See also G. W. Scherr and R. E. Knight, Rev. Mod. Phys. 35, 436 (1963).
5. R. M. Delves, Nuc. Phys. 41, 497 (1963).
6. For the ground state we know $\tilde{E}^{(0)} + \sqrt{\tilde{E}^{(1)}} + \dots \geq E^{(0)} + \sqrt{E^{(1)}} + \dots$.
If now our trial function is accurate to order $\sqrt[n]{}$ then the left hand side will be accurate to order $\sqrt[2n+1]{}$ whence we will have
$$\sqrt[2n+2]{\tilde{E}^{(2n+2)}} + \dots \geq \sqrt[2n+2]{E^{(2n+2)}} + \dots$$

which implies $\tilde{E}^{(2n+2)} \geq E^{(2n+2)}$. In short we will have minimal principles only for even orders, and then only if we know the lower orders exactly.
7. O. Sinanoglu, J. Chem. Phys. 34, 1237 (1961).
G. W. Scherr, and R. E. Knight, Rev. Mod. Phys. 35, 436 (1963).
8. As has already been remarked in Chapter III, in many applications \tilde{J}_H has been used as an approximation to $-E^{(2)}$. As is clear from our discussion, and as most authors have realized and mentioned, this is not a variational approximation unless $\tilde{\psi}^{(0)} = \psi^{(0)}$.

9. S. I. Vetchinkin, Optics and Spectroscopy, 14, 169 (1963); and F. Villars in Proc. of the Int. School of Physcs. Enrico Fermi, Course XXIII, Nuclear Physics (Academic Press, New York, 1963).
10. M. Cohen and A. Dalgarno, Rev. Mod. Phys., 35, 506 (1963).
D. Layzer, Phys. Rev. 132, 735 (1963).
11. C. S. Sharma and C. A. Coulson, Proc. Phys. Soc. 80, 81 (1962).
12. A different theorem (with much weaker conditions) is stated by R. E. Knight and C. W. Scherr, Rev. Mod. Phys. 35, 431 (1963), Appendix II, but this in fact is the theorem which they prove.
13. For the ordinary variational principle $\delta J = 0$ and hence by implication for $\delta J^{(0)} = \delta J^{(1)} = \dots = 0$ the problem has already been discussed by A. C. Hurley, Proc. Roy. Soc. A226, 179 (1954) who shows that a sufficient condition is that the trial function not depend explicitly on the parameter in question.
14. They are of course trivially equal if \mathcal{H}_0 , and hence $\psi^{(0)}$, is independent of σ . For a further discussion of this case see R. Yaris, J. Chem. Phys. 39, 863 (1963). They are also of course equal if $\tilde{\psi}^{(1)} = \psi^{(1)}$.

* Here and in what follows η in a general way, symbolizes a small quantity. It need not have the same value in different variations.

Note also that since J is homogeneous the variations

$$\delta \tilde{\psi}_p = \eta \tilde{\psi}_p \text{ and } \delta \tilde{\psi}_q = \eta \tilde{\psi}_q \text{ are, in effect, always allowed.}$$

Thus all conditions are to be understood as being modulo such variations.

VI. THE FOPIM AND FOP-VIM FAST CONVERGING ITERATION PROCEDURES

As we have discussed in Chapter II, given the wave function correct to $O(\lambda^n)$ we can compute the energy correct to $O(\lambda^{2n+1})$. Although this is "better" than computing the energy only to $O(\lambda^n)$, this process is not rapidly converging since the energy is only improved two orders in λ for each additional order in the wave function. In this Chapter we present two procedures which are much more rapidly convergent: the First Order Perturbation Method or FOPIM, and a slightly improved version called First Order Perturbation-Variation Iterative Method or FOP-VIM. In FOPIM, the wave function through the first order is used as the zeroeth order function in the calculation of an improved first-order function. The perturbation potential for this new calculation is proportional to λ^2 . This process can be iterated and each time the new perturbation potential is proportional to the square of the previous perturbation potential.¹ Thus, after n iterations, the energy is given accurately up to terms of the order of λ raised to the $(2)^{n+1}$ power. For example, after 5 iterations the energy is accurate up to terms of the order of λ^{64} .

In FOP-VIM, following Dalgarno and Stewart (1961), the perturbed wave function is taken to be the variationally best linear combination of the zeroeth order and Rayleigh-Schrodinger first order functions. This perturbed-variational function is then taken to be the zeroeth-order wave function for the calculation of an improved perturbed-variational function. Whereas FOP-VIM may have only a modest advantage over FOPIM for non-degenerate energy levels, we show in Chapter VII that the basic notion of FOP-VIM is very useful for degenerate or almost degenerate

energy levels.

We start with a trial wave function ψ which satisfies the Schrodinger equation $H_0 \psi = \epsilon \psi$. The Hamiltonian for the perturbed system is $H = H_0 + \lambda V$. The first order wave function is $\psi^{(1)} = F\psi$. The expectation value of the energy corresponding to the original trial function is $\epsilon(0)$ as given by Eq. (2.15). The perturbed normalized wave function through the first order is $\Psi(1)$ as given by Eqs. (2.13) and (2.14). The expectation value of the energy corresponding to $\Psi(1)$ is given by $\epsilon(1)$ of Eq. (2.16). Everything up to this point has been discussed in the preceding chapters. The following treatment is somewhat novel.

FOPIM

The wave function $\Psi(1)$ satisfies the Schrodinger equation

$$H(1)\Psi(1) = \epsilon(0)\Psi(1) \quad (1)$$

where

$$H(1) = H_0 + \frac{\lambda V + \lambda^2 \epsilon^{(1)} F}{1 + \lambda F} \quad (2)$$

The Hamiltonian for the perturbed system is then

$$H = H(1) + \lambda^2 V(1) \quad (3)$$

where

$$V(1) = \frac{F(V - \epsilon^{(1)})}{1 + \lambda F} \quad (4)$$

Now we consider the new perturbation problem in which $\Psi(1)$ serves as the zeroeth order wave function and $\lambda^2 V(1)$ is the perturbation

potential. Note the λ^2 plays the same role in the new problem as λ plays in the original problem. The new first order perturbed wave function is

$$\psi^{(1)}(1) = F(1)\Psi(1) \quad (5)$$

where the function $F(1)$ is determined by

$$[F(1), H]\Psi(1) = (V(1) - \epsilon^{(1)}(1))\Psi(1) \quad (6)$$

The methods of solving equations of this form are discussed in Chapter III. The specification of the $F(1)$ is completed by the requirement that

$$\langle \Psi(1), F(1)\Psi(1) \rangle = 0 \quad (7)$$

square
The/norm of $\psi^{(1)}(1)$ is designated as $S(1)$,

$$\langle F(1)\Psi(1), F(1)\Psi(1) \rangle = S(1) \quad (8)$$

Using the first, second, and third order perturbation equations for the new perturbation problem,

$$\epsilon^{(1)}(1) = \langle \Psi(1), V(1)\Psi(1) \rangle = [\mathcal{E}(1) - \mathcal{E}(0)]/\lambda^2 \quad (9)$$

$$\epsilon^{(2)}(1) = \langle \Psi(1), V(1)F(1)\Psi(1) \rangle \quad (10)$$

$$\epsilon^{(3)}(1) = \langle F(1)\Psi(1), V(1)F(1)\Psi(1) \rangle - \epsilon^{(1)}(1)S(1) \quad (11)$$

We can now define the second iterated normalized wave function

$$\psi(2) = [1 + \lambda^4 S(1)]^{-\frac{1}{2}} (1 + \lambda^2 F(1))\Psi(1) \quad (12)$$

The expectation value of the energy corresponding to $\psi(2)$ is

$$E(2) = \langle \psi(2), H\psi(2) \rangle = E(1) + \frac{\lambda^4 \epsilon^{(2)}(1) + \lambda^6 \epsilon^{(3)}(1)}{1 + \lambda^4 S(1)} \quad (13)$$

For exactly the same reasons the $E(1)$ is accurate up to terms of the order of λ^4 , it is apparent that $E(2)$ is accurate up to terms of the order of λ^8 .

The function $\psi(2)$ can now be used as the zeroth order wave function in a new perturbation calculation. Indeed, after $(n+1)$ iterations we have the normalized wave function

$$\psi(n+1) = [1 + \lambda^{2a} S(n)]^{-\frac{1}{2}} (1 + \lambda^a F(n)) \psi(n) \quad (14)$$

Here, because of printing difficulties, we use the notation $a = 2^n$.
square
The $S(n)$ is the/norm of $\psi^{(1)}(n)$,

$$\langle F(n)\psi(n), F(n)\psi(n) \rangle = S(n) \quad (15)$$

The function $F(n)$ satisfies the inhomogenous partial differential equation

$$[F(n), H] \psi(n) = [V(n) - \epsilon^{(1)}(n)] \psi(n) \quad (16)$$

together with the requirement that

$$\langle \psi(n), F(n)\psi(n) \rangle = 0 \quad (17)$$

From the first, second, and third order perturbation equations for this perturbation problem it follows that

$$\epsilon^{(1)}(n) = \langle \psi(n), V(n)\psi(n) \rangle = [E(n) - E(n-1)] / \lambda^a \quad (18)$$

$$\epsilon^{(2)}(n) = \langle \psi(n), V(n)F(n)\psi(n) \rangle \quad (19)$$

$$\epsilon^{(3)}(n) = \langle F(n)\psi(n), V(n)F(n)\psi(n) \rangle - \epsilon^{(1)}(n)S(n) \quad (20)$$

The expectation value of the energy corresponding to $\psi(n+1)$ is

$$E(n+1) = \langle \psi(n+1), H\psi(n+1) \rangle = E(n) + \frac{\lambda^{2a} \epsilon^{(2)}(n) + \lambda^{3a} \epsilon^{(3)}(n)}{1 + \lambda^{2a} S(n)} \quad (21)$$

Since λ^a in the present problem plays the same role as λ in the original problem, it follows that since $\epsilon^{(1)}$ is accurate up to terms of the order of λ^4 , the value of $E(n+1)$ is accurate up to terms of the order of λ^{4a} .

The function $\psi(n+1)$ satisfies the Schrodinger equation

$$H(n+1)\psi(n+1) = E(n)\psi(n+1) \quad (22)$$

where

$$H(n+1) = H(n) + \frac{\lambda^a V(n) + \lambda^{2a} \epsilon^{(1)}(n) F(n)}{1 + \lambda^a F(n)} \quad (23)$$

The Hamiltonian for the perturbed system is then

$$H = H(n+1) + \lambda^{2a} V(n+1) \quad (24)$$

where

$$V(n+1) = \frac{F(n) [V(n) - \epsilon^{(1)}(n)]}{1 + \lambda^a F(n)} \quad (25)$$

The iteration scheme can then be continued through the next step.

FOP-VIM

The perturbed wave functions can be improved still further with very little effort.² This involves the coupling of the perturbation and variational techniques. Consider in place of $\Psi(1)$, the normalized function

$$\chi(1) = [1 + \lambda^2 \alpha^2 S]^{-\frac{1}{2}} (1 + \lambda \alpha F) \psi \quad (26)$$

Here $S = \langle F\psi, F\psi \rangle$ and α is a parameter varied to make stationary the energy $E'(1) = \langle \chi(1), H\chi(1) \rangle$. For this optimum value of α , the value of $E'(1)$ is given by the secular equation

$$\begin{vmatrix} E(0) - E'(1) & \lambda^2 \epsilon^{(2)} \\ \lambda^2 \epsilon^{(2)} & -\lambda^2 \epsilon^{(2)} + \lambda^3 \epsilon^{(3)} + \lambda^2 S [E(0) - E'(1)] \end{vmatrix} = 0 \quad (27)$$

Of the two roots to the secular equation, the only root which has physical significance for our problem corresponds to $E'(1)$ approaching $E(0)$ as λ approaches zero. Thus,

$$E'(1) = E(0) + \left[\frac{\epsilon^{(2)} - \lambda \epsilon^{(3)}}{2S} \right] \left[-1 + \left\{ 1 + \frac{4\lambda^2 S [\epsilon^{(2)}]^2}{[\epsilon^{(2)} - \lambda \epsilon^{(3)}]^2} \right\}^{-\frac{1}{2}} \right] \quad (28)$$

The improvement in the energy obtained by this variation of α is given by the expansion in powers of λ ,

$$E'(1) - E(1) = \lambda^4 \frac{[\epsilon^{(3)}]^2}{\epsilon^{(2)}} + \lambda^5 \left[\frac{[\epsilon^{(3)}]^3}{[\epsilon^{(2)}]^2} - 2S \epsilon^{(3)} \right] + \dots \quad (29)$$

The optimum value of α is $[E'(1) - E(0)][\lambda^2 \epsilon^{(2)}]^{-1}$, or expanding in powers of λ ,

$$\alpha = 1 + \lambda \frac{\epsilon^{(3)}}{\epsilon^{(2)}} + \lambda^2 \left[\left(\frac{\epsilon^{(3)}}{\epsilon^{(2)}} \right)^2 - s \right] + \dots \quad (30)$$

Clearly, in much the same manner as in FOPIM, the $\chi(1)$ can be used as the zeroeth order wave function to generate a new first order function. The $\chi(1)$ satisfies the Schrodinger equation

$$h(1)\chi(1) = E(0)\chi(1) \quad (31)$$

where

$$h(1) = H_0 + \frac{\lambda \alpha V + \lambda(1-\alpha)\epsilon^{(1)} + \lambda^2 \alpha \epsilon^{(1)}_F}{1 + \lambda \alpha F} \quad (32)$$

The Hamiltonian for the perturbed system is then $H = h(1) + \lambda^2 V'(1)$ where

$$V'(1) = [\lambda^{-1}(1-\alpha) + \alpha F] [V - \epsilon^{(1)}] [1 + \lambda \alpha F]^{-1} \quad (33)$$

From Eq. (31) it follows that $\lambda^{-1}(1-\alpha)$ is zeroeth order in λ .

FOOTNOTES TO CHAPTER VI

1. J. O. Hirschfelder, J. Chem. Phys. 39, 2009 (1963). There is a variety of other iterative procedures that have been proposed to take advantage of this rapid convergence. The most noteworthy of these are L. Biedenharn and J. M. Blatt, Phys. Rev. 93, 230 (1954); T. Kikuta, Prog. Theo. Phys. 12, 10 (1954) and 14, 453 (1955); and R. A. Sack, University of Wisconsin Theoretical Chemistry Institute Report No. 30 (1963). In addition, P. O. Lowdin, J. Molec. Spect. 10, 12 (1963) has developed an iterative perturbation-variation procedure for solving the perturbed Schrodinger equation by partitioning the secular equation.
2. A. Dalgarno and A. L. Stewart, Proc. Phys. Soc. (London) 77, 467 (1961). A similar procedure has been used in connection with the Brillouin-Wigner perturbation method by P. Goldhammer and E. Feenberg, Phys. Rev. 101, 1233 (1956); and by R. C. Young, L. C. Biedenharn, and E. Feenberg, Phys. Rev. 106, 1151 (1957).

VII. DEGENERATE OR ALMOST DEGENERATE ENERGY LEVEL PERTURBATIONS

Non-degenerate perturbation theory becomes inapplicable when two or more energy levels lie close together and interact strongly under the influence of the perturbation. An extension of the FOP-VIM analysis of Chapter VI suffices to show the explicit effect of some quantum state k on the q -th state under consideration. Let us form a trial wave function $\Phi_q(1)$ as a linear combination of ψ_k and $\chi_q(1)$, as defined by Eq. (6.26),

$$\Phi_q(1) = \chi_q(1) + \lambda^2 C \psi_k \quad (1)$$

The constant C can be adjusted so as to optimize the energy

$$\tilde{E}_q = \langle \Phi_q(1), H \Phi_q(1) \rangle / \langle \Phi_q(1), \Phi_q(1) \rangle \quad (2)$$

subject to the condition that, as λ approaches zero, \tilde{E}_q approaches ϵ_q . In this Section it is convenient to use the notation:

$(X)_{ij} = \langle \psi_i, X \psi_j \rangle$. Since $\psi_q^{(1)} = F_q \psi_q$, it follows from Eq. (2.17) that

$$(F_q)_{kq} = (V)_{kq} [\epsilon_q - \epsilon_k]^{-1} \quad (3)$$

The constant C corresponding to the physically significant root of the secular equation is

$$C = [\epsilon_q - \epsilon_k]^{-2} [-\epsilon_q^{(1)} - \{\epsilon_q - \epsilon_k\} \frac{\epsilon_q^{(3)}}{\epsilon_q^{(2)}} + (F_q V)_{qk}] + \lambda [\dots] + \dots \quad (4)$$

Expanding the energy in powers of λ ,

$$\begin{aligned} \tilde{E}_q = E_q^{(1)} + \lambda^4 [\epsilon_q - \epsilon_k]^{-3} [-\epsilon_q^{(1)} - \{\epsilon_q - \epsilon_k\} \frac{\epsilon_q^{(3)}}{\epsilon_q^{(2)}} + (F_q V)_{qk}] \\ + \lambda^5 [\dots] + \dots \end{aligned} \quad (5)$$

Here $E'_q(1)$ is given by Eq. (6.28). Since $E'_q(1)$ is accurate through terms in λ^3 it is not surprising that the first corrections to $E'_q(1)$ should be proportional to λ^4 . From Eq. (5) it follows that $\tilde{E}_q > E'_q(1)$ if $\epsilon_q > \epsilon_k$ and $\tilde{E}_q < E'_q(1)$ if $\epsilon_q < \epsilon_k$. For the ground state or the lowest energy state of a given symmetry, the energy is lowered by the interaction with every other state. If $|\tilde{E}_q - E'_q(1)| > |E'_q(1) - E_q(0)|$, the expansion of \tilde{E}_q in powers of λ (as given by Eq. (5)) is not longer possible and the non-degenerate perturbation theory becomes inapplicable.

DE-FOP-VIM

The usual Rayleigh-Schrodinger method can be exceedingly complex and difficult to apply to perturbation problems involving degenerate or almost-degenerate energy levels (see Dalgarno, 1961). Greater accuracy and far greater simplicity can be obtained by the following procedure: DE-FOP-VIM, which is FOP-VIM generalized to apply to the degenerate and almost degenerate energy levels (Hirschfelder, 1963).

We may wish to consider the effect of a perturbation either on a particular degenerate energy level or the effect on a tightly packed group of degenerate or almost degenerate energy levels. Since the states corresponding to different energies may interact under the influence of a perturbation, the larger the number of interacting states which are explicitly considered, the greater is the accuracy of the calculations.

The first step in the DE-FOP-VIM is the determination of the "starting" wave functions.¹ For each of the n_k degenerate energy states ϵ_k under consideration, we are given a complete set of

linearly independent eigenfunctions $\phi_{k1}, \dots, \phi_{kn_k}$ for the unperturbed Hamiltonian H_0 . Thus, $H_0 \phi_{k\beta} = \epsilon_k \phi_{k\beta}$. The Hamiltonian for the perturbed system is $H_0 + \lambda V$. The "starting" functions are

$$\psi_{k\alpha} = \sum_{\beta=1}^{n_k} a_{\alpha\beta} \phi_{k\beta} \quad (6)$$

The constants $a_{\alpha\beta}$ are chosen to diagonalize the perturbation matrix so that

$$\langle \psi_{k\alpha}, V \psi_{k\beta} \rangle = \delta_{\alpha\beta} V_{k\alpha; k\alpha} \quad (7)$$

This automatically makes the different $\psi_{k\alpha}$ orthogonal,

$$\langle \psi_{k\alpha}, \psi_{k\beta} \rangle = 0, \quad \alpha \neq \beta \quad (8)$$

If each of the eigenvalues $V_{k\alpha; k\alpha}$ of the matrix is different, the degeneracy is removed in the first order of the perturbation and the choice of the $a_{\alpha\beta}$ is determined to within a phase factor. If, on the other hand, the eigenvalues for the energy matrix are not all different, the $a_{\alpha\beta}$ are not completely specified. For our purposes, in contrast to the usual Rayleigh-Schrodinger treatment, any choice of the $a_{\alpha\beta}$ compatible with Eq. (7) is sufficient. This is one of the principle simplifications of the DE-FOP-VIM.

Corresponding to a "starting" wave function $\psi_{k\alpha}$ we can define a function $F_{k\alpha}$ such that $F_{k\alpha} \psi_{k\alpha}$ satisfies the first order equation,

$$(H_0 - \epsilon_k) F_{k\alpha} \psi_{k\alpha} + [V - V_{k\alpha; k\alpha}] \psi_{k\alpha} = 0 \quad (9)$$

The methods for solving equations of this type are discussed in Chapter III. Any solution of Eq. (9) suffices provided that $F_{k\alpha} \psi_{k\alpha}$ satisfies

the usual boundary, continuity, and integrability conditions required of a bound state wave function. The fact that

$$G_{k\alpha} = F_{k\alpha} + \sum_{\beta=1}^{n_k} b_{\alpha\beta} \psi_{k\beta} / \psi_{k\alpha} \quad (10)$$

for arbitrary constants $b_{\alpha\beta}$ is also a solution to Eq. (9) does not cause us any concern.

Now we are ready to consider the effect of the perturbation on a set of $n = \sum_k n_k$ energy levels ϵ_k . We fix our attention on those perturbed wave functions Ψ_q whose energies E_q approach the values of the ϵ_k in the limit as λ approaches zero. The first approximation to Ψ_q can then be written in the form

$$\psi_q(1) = \sum_k \sum_{\alpha=1}^{n_k} [J_{q;k\alpha} + \lambda K_{q;k\alpha} F_{k\alpha}] \psi_{k\alpha} \quad (11)$$

The constants $J_{q;k\alpha}$ and $K_{q;k\alpha}$ are chosen so as to normalize $\psi_q(1)$ and optimize the energy $E_q(1) = \langle \psi_q(1), H \psi_q(1) \rangle$ subject to the condition that $E_q(1)$ approaches the value of one of the ϵ_k in the limit as λ approaches zero. The optimum values of $E_q(1)$ are given as n of the $2n$ roots of the secular equation

$$\begin{vmatrix} A & B \\ B^\dagger & C \end{vmatrix} = 0 \quad (12)$$

where A , B , and C are the n by n dimensional submatrices with the elements

$$A_{k\alpha;k'\beta} = \langle \psi_{k\alpha}, H \psi_{k'\beta} \rangle - E(1) \langle \psi_{k\alpha}, \psi_{k'\beta} \rangle \quad (13)$$

$$B_{k\alpha;k'\beta} = \langle \psi_{k\alpha}, H F_{k'\beta} \psi_{k'\beta} \rangle - E(1) \langle \psi_{k\alpha}, F_{k'\beta} \psi_{k'\beta} \rangle \quad (14)$$

$$C_{k\alpha;k'\beta} = \langle F_{k\alpha} \psi_{k\alpha}, H F_{k'\beta} \psi_{k'\beta} \rangle - E(1) \langle F_{k\alpha} \psi_{k\alpha}, F_{k'\beta} \psi_{k'\beta} \rangle \quad (15)$$

The Lowdin (1963) partitioning technique is particularly well suited to the solution of such secular equations. The values of the $E_q(1)$ obtained in this manner should be accurate through terms of $O(\lambda^3)$.

The $\psi_q(1)$ can themselves be used as the zeroeth order wave functions for a new perturbation calculation in much the same manner as in the FOPIM or FOP-VIM procedures.

FOOTNOTES TO CHAPTER VII

1. The advantage of determining the "correct zeroeth order" functions $\psi_{k\alpha}$ is that the first order perturbation energies are obtained at an early stage in the calculations. If, however, one wishes to avoid this step, the $\phi_{k\alpha}$ can be used in place of the $\psi_{k\alpha}$ in Eq. (10) and beyond provided that Eq. (9) is replaced by

$$(H_0 - E_k) \phi_{k\alpha} + V \phi_{k\alpha} - \sum_{\beta=1} V_{k\beta; k\alpha} \phi_{k\beta} = 0 \quad (9')$$

where $V_{k\beta; k\alpha} = \langle \phi_{k\beta}, V \phi_{k\alpha} \rangle$. Equation (9') follows from the multiplication of Eq. (9) by $a_{\beta\alpha}^{-1}$, summing over α , and making use of Eq. (6).

VIII. OFF-DIAGONAL MATRIX ELEMENTS

As we have seen in Chapter IV, the Dalgarno Interchange theorem is often of decisive importance in making practical the calculation of the first order corrections to the expectation values of one-particle operators. In this chapter we show that a similar theorem holds for off-diagonal matrix elements.¹

A. General Formulation

In order to compute the matrix element $\langle \Psi_p, W \Psi_q \rangle$ for a Hermitian operator W between the eigenfunctions Ψ_p and Ψ_q of two Hamiltonians H_p and H_q (possibly, but not necessarily, identical), we consider the equations²

$$H_p \Phi_p + \mu s W \Phi_q = \epsilon_p \Phi_p, \quad (1)$$

$$H_q \Phi_q + \mu s^* W \Phi_p = \epsilon_q \Phi_q, \quad (2)$$

where s is an arbitrary complex number.

Introducing a 2x2 matrix notation we can combine (1) and (2) into

$$\underline{H} \underline{\Phi} = \underline{\epsilon} \underline{\Phi} \quad (3)$$

where

$$\underline{H} = \begin{pmatrix} H_p & \mu s W \\ \mu s^* W & H_q \end{pmatrix}, \quad \underline{\epsilon} = \begin{pmatrix} \epsilon_p & 0 \\ 0 & \epsilon_q \end{pmatrix}, \quad \underline{\Phi} = \begin{pmatrix} \Phi_p \\ \Phi_q \end{pmatrix}$$

Equation (3) is now of standard Hamiltonian form, and thus, with suitable notational changes, we can take over many of the earlier results in this article. In particular, if we write $H_p = H_{po} + \lambda V_p$,

$$H_q = H_{q0} + \lambda V_q, \quad \underline{\Psi} = \underline{H}_0 + \lambda \underline{V} + \mu \underline{W}, \quad \underline{\Phi} = \underline{\Psi}^{(0,0)} + \lambda \underline{\Psi}^{(1,0)} + \mu \underline{\Psi}^{(0,1)} + \dots$$

and $\underline{\Xi} = \underline{\Xi}^{(0,0)} + \lambda \underline{\Xi}^{(0,1)} + \mu \underline{\Xi}^{(1,0)} + \dots$, we have the identity

$$\langle \underline{\Psi}^{(1,0)}, \underline{W} \underline{\Psi}^{(0,0)} \rangle + \langle \underline{\Psi}^{(0,0)}, \underline{W} \underline{\Psi}^{(1,0)} \rangle = \langle \underline{\Psi}^{(0,1)}, \underline{V} \underline{\Psi}^{(0,0)} \rangle + \langle \underline{\Psi}^{(0,0)}, \underline{V} \underline{\Psi}^{(0,1)} \rangle \quad (4)$$

where

$$(\underline{H}_0 - \underline{\Xi}^{(0,0)}) \underline{\Psi}^{(0,0)} = 0 \quad (5)$$

$$(\underline{H}_0 - \underline{\Xi}^{(0,0)}) \underline{\Psi}^{(1,0)} = -(\underline{V} - \underline{\Xi}^{(1,0)}) \underline{\Psi}^{(0,0)} \quad (6)$$

$$(\underline{H}_0 - \underline{\Xi}^{(0,0)}) \underline{\Psi}^{(0,1)} = -(\underline{W} - \underline{\Xi}^{(0,1)}) \underline{\Psi}^{(0,0)} \quad (7)$$

with

$$\underline{\Xi}^{(1,0)} = \begin{pmatrix} \langle \underline{\Psi}_p^{(0,0)}, \underline{V} \underline{\Psi}_p^{(0,0)} \rangle & 0 \\ 0 & \langle \underline{\Psi}_q^{(0,0)}, \underline{V} \underline{\Psi}_q^{(0,0)} \rangle \end{pmatrix};$$

$$\underline{\Xi}^{(0,1)} = \begin{pmatrix} s \langle \underline{\Psi}_p^{(0,0)}, \underline{W} \underline{\Psi}_q^{(0,0)} \rangle & 0 \\ 0 & s^* \langle \underline{\Psi}_q^{(0,0)}, \underline{W} \underline{\Psi}_p^{(0,0)} \rangle \end{pmatrix}.$$

and where we have normalized in such a way that

$$\langle \underline{\Psi}_p^{(0,0)}, \underline{\Psi}_p^{(1,0)} \rangle = \langle \underline{\Psi}_p^{(0,0)}, \underline{\Psi}_p^{(0,1)} \rangle = \langle \underline{\Psi}_q^{(0,0)}, \underline{\Psi}_q^{(1,0)} \rangle = \langle \underline{\Psi}_q^{(0,0)}, \underline{\Psi}_q^{(0,1)} \rangle = 0$$

For $s = 1$, the left hand side of Eq. (4) is simply twice the first order (in λ) correction for the matrix element $\langle \underline{\Psi}_p, \underline{W} \underline{\Psi}_q \rangle$, and for $s = -i$ it is twice the imaginary part since

$$\underline{\Psi}_p = \underline{\Psi}_p^{(0,0)} + \lambda \underline{\Psi}_p^{(1,0)} + \dots \quad \text{and} \quad \underline{\Psi}_q = \underline{\Psi}_q^{(0,0)} + \lambda \underline{\Psi}_q^{(1,0)} + \dots$$

The right hand side of Eq. (4) then informs us that we can equally

well evaluate these corrections by solving Eq. (7) rather than Eq. (6), i.e. we have an interchange theorem. We will discuss the solution of Eq. (7) in Section C below.

B. Variation Principles

Equation (3) can of course be derived from the variation principle $\delta J = 0$ where

$$J = \langle \tilde{\Phi}, (\underline{H} - \underline{E}) \tilde{\Phi} \rangle$$

We can carry over to the off-diagonal case much of our earlier discussion (Chapter V) of connections between perturbation theory and the variational principle. In particular, writing $\underline{H} = \underline{H} + \mu \underline{W}$, $\tilde{\Phi} = \tilde{\Phi}^{(0)} + \mu \tilde{\Phi}^{(1)} + \dots$ and $\underline{E} = \underline{E}^{(0)} + \mu \underline{E}^{(1)} + \dots$, the variational principle $\delta J^{(1)} = 0$, $J^{(1)} = 0$, $J^{(0)} = 0$ with

$$J^{(0)} \equiv \langle \tilde{\Phi}^{(0)}, (\underline{H} - \underline{E}^{(0)}) \tilde{\Phi}^{(0)} \rangle$$

$$J^{(1)} \equiv \langle \tilde{\Phi}^{(0)}, (\underline{W} - \underline{E}^{(1)}) \tilde{\Phi}^{(0)} \rangle + \langle \tilde{\Phi}^{(0)}, (\underline{H} - \underline{E}^{(0)}) \tilde{\Phi}^{(1)} \rangle + \langle \tilde{\Phi}^{(1)}, (\underline{H} - \underline{E}^{(0)}) \tilde{\Phi}^{(0)} \rangle$$

gives us, for $s = 1$, a variation principle for twice the real part of the desired matrix element and for $s = -i$ with a variation principle for twice the imaginary part. These are, with appropriate notational changes, exactly the variational principles given by Delves² and have the advantages and disadvantages, already discussed in Chapter V, of his variational principle for diagonal matrix elements. If further, in analogy to the Schwartz procedure for diagonal matrix

elements, for a given $\tilde{\Phi}_{\underline{p}}^{(0)}$ we constrain $\tilde{\Phi}_{\underline{p}}^{(1)}$ so that $\tilde{\Phi}_{\underline{p}}^{(0)} - \Phi_{\underline{p}}^{(0)}$ and $\tilde{\Phi}_{\underline{p}}^{(1)} - \Phi_{\underline{p}}^{(1)}$ are of the same order (this is the approach in effect followed by Borowitz¹) then, of course, as in the diagonal case, one arrives at results quite equivalent to those of our double perturbation approach.

C. Separation of Equations for One-Electron W

For W a one particle operator, we will show that our problem reduces to the solution of one electron equations in the special case for which $\psi_p^{(0,0)}$ and $\psi_q^{(0,0)}$ are single determinants. The method is obviously extendable to the case where only one of the $\psi^{(0,0)}$ is a single determinant. The case where neither $\psi_p^{(0,0)}$ nor $\psi_q^{(0,0)}$ is a single determinant remains to be investigated.

Let us consider then $\psi_p^{(0,0)} = \star \phi_1(1) \phi_2(2) \dots \phi_N(N)$ and $\psi_q^{(0,0)} = \star \bar{\phi}_1(1) \phi_2(2) \dots \phi_N(N)$ where \star is the antisymmetrization operator and where we assume $\psi_p^{(0,0)}$ and $\psi_q^{(0,0)}$ differ by only a single spin orbital³ (the case where they differ by more than one is much simpler in that $\xi^{(0,1)} \equiv 0$). Since all the operators in (7) are symmetric in the particles, we clearly may write $\xi^{(0,1)} = \star \chi$ whence, singling out one component of Eq. (7) in detail, we have

$$(h_1 + h_2 + \dots - e_1 - e_2 - \dots) \chi_p = -s(w_1 + w_2 + \dots) \bar{\phi}_1(1) \phi_2(2) \dots + s \langle \phi_1, w_1 \bar{\phi}_1 \rangle \phi_1(1) \phi_2(2) \dots$$

Here $H_{op} = \sum_i h_i$, $E_p^{(0,0)} = \sum_i e_i$. Also we need $h_1 \bar{\phi}_1(1) = \bar{e}_1 \bar{\phi}_1(1)$

(we will assume, for simplicity, that $H_{op} = H_{oq}$). We now observe

that we can write $\chi_p = \chi_{p1} + \chi_{p2}$ where

$$(h_1 + h_2 + \dots - e_1 - e_2 - \dots) \chi_{p1} = -sw_1 \bar{\varphi}_1(1) \varphi_2(2) \dots$$

$$+ s \langle \varphi_1, w_1 \bar{\varphi}_1 \rangle \varphi_1(1) \varphi_2(2) \dots$$

and

$$(h_1 + h_2 + \dots - e_1 - e_2 - \dots) \chi_{p2} = -s(w_2 + w_3 + \dots) \bar{\varphi}_1(1) \varphi_2(2) \dots$$

the essential point being that the right hand sides of each of the last two equations are orthogonal to the solution of the homogeneous equation, and therefore the equations are consistent.

The reduction to one electron equations is now immediate, namely

$$\chi_{p1} = \Delta_1(1) \varphi_2(2) \dots \varphi_N(N)$$

$$\chi_{p2} = \bar{\varphi}_1(1) \Delta_2(2) \varphi_3(3) \dots \varphi_N(N) + \bar{\varphi}_1(1) \varphi_2(2) \Delta_3(3) \dots + \dots$$

where

$$(h_1 - e_1) \Delta_1 = -sw_1 \bar{\varphi}_1 + s \langle \varphi_1, w_1 \bar{\varphi}_1 \rangle \varphi_1$$

$$(h_2 + \bar{e}_1 - e_2) \Delta_2 = -sw_2 \varphi_2$$

.....

Further reduction depends on the particular form of W . Note also that if V is a one particle operator (as it is within the Hartree-Fock formalism), then χ_{p2} does not contribute to the right hand side of Eq. (4) and our theory, as it must, becomes identical to that of Cohen and Dalgarno.¹

Finally we remark that, as with the diagonal case, variational

procedures are also available to approximate the solution of Eq. (7), the stationary expression being just the Hylleraas form

$$J_H = \langle \tilde{\psi}_{\underline{0}}^{(0,1)}, (H_0 - \underline{\epsilon}^{(0,0)}) \tilde{\psi}_{\underline{0}}^{(0,1)} \rangle + \langle \tilde{\psi}_{\underline{0}}^{(0,1)}, (W - \underline{\epsilon}^{(0,1)}) \tilde{\psi}_{\underline{0}}^{(0,0)} \rangle + \langle \tilde{\psi}_{\underline{0}}^{(0,0)}, (W - \underline{\epsilon}^{(0,1)}) \tilde{\psi}_{\underline{0}}^{(0,1)} \rangle .$$

FOOTNOTES TO CHAPTER VIII

1. Within the Hartree-Fock formalism this has been shown by M. Cohen, and A. Dalgarno, Proc. Roy. Soc. (London) A275, 492 (1963). The general result to be derived in this Chapter has also been derived by S. Borowitz (private communication).
2. Such equations have first been discussed by L. M. Delves, Nuc. Phys. 45, 313 (1963).
3. We ignore possible complications due to degeneracy. It often is the case that W is "diagonal" in the sense that given $\psi_p^{(0,0)}$ then $\langle \psi_q^{(0,0)}, W \psi_p^{(0,0)} \rangle$ differs from zero only for one $\psi_q^{(0,0)}$ of the degenerate set, and similarly given $\psi_q^{(0,0)}$ a unique $\psi_p^{(0,0)}$ is singled out. We also assume $\langle \bar{\phi}_1, \phi_1 \rangle = 0$.